

Uniwersytet Warszawski
Wydział Fizyki

Maciej Jagielski

Rozprawa doktorska pt.:

Zastosowanie nieliniowego równania Langevina,
równania Fokkera–Plancka oraz modeli błędzeń
losowych do opisu dochodów gospodarstw domowych
Polski i Unii Europejskiej

Promotor:

prof. dr hab. Ryszard Kutner

Zakład Fizyki Biomedycznej

Instytut Fizyki Doświadczalnej WF UW

Warszawa, lipiec 2013

Univeristy of Warsaw
Faculty of Physics

Maciej Jagielski

Title of dissertation:

Application of nonlinear Langevin equation,
Fokker–Planck equation, and random walk models to
modelling of household incomes in Poland and the
European Union

Dissertation directed by:
professor Ryszard Kutner
Section of Biomedical Physics
Institute of Experimental Physics

Warsaw, July 2013

PODZIĘKOWANIA

Chciałem złożyć serdeczne podziękowania profesor Tizianie Di Matteo z Wydziału Matematyki King's College London oraz profesorowi Victorowi M. Yakovenko z Wydziału Fizyki Uniwersytetu Maryland za szereg cennych i stymulujących uwag.

Dziękuję również mojemu promotorowi profesorowi Ryszardowi Kutnerowi za poświęcony czas i cierpliwe sprawdzanie moich obliczeń.

Pragnę wyrazić głęboką wdzięczność Panu Markowi Pęczkowskiemu z Wydziału Nauk Ekonomicznych Uniwersytetu Warszawskiego za pomoc przy opracowaniu baz danych.

Szczególne podziękowania kieruję do mojej Rodziny i Przyjaciół, za ich miłość, pomoc, cierpliwość, wiarę we mnie oraz poczucie humoru, które podnosiło mnie na duchu w trudnych momentach pracy nad doktoratem.

STRESZCZENIE

Niniejsza rozprawa doktorska stanowi oryginalne zastosowanie wybranych kanonicznych modeli używanych w ramach fizyki statystycznej do badania własności układów leżących poza zainteresowaniem tradycyjnie rozumianej fizyki. Mam tu na myśli obszary badawcze będące tradycyjnie domeną ekonomii oraz socjologii a dzisiaj stanowiące już przedmiot intensywnych i ekstensywnych badań ekono- i socjofizyki.

Dokładniej rzecz biorąc, tematyka niniejszej rozprawy doktorskiej dotyczy dynamiki dochodów gospodarstw domowych społeczeństw Polski oraz Unii Europejskiej w okresach kryzysów i poza nimi. Badania przeprowadziłem zarówno w ramach dynamiki stochastycznej Langevina i równania Fokkera–Plancka, jak też wybranych modeli błędów losowych, gdzie dochód gospodarstw domowych jest traktowany jak zależna od czasu zmienna losowa.

W rozprawie uzyskałem wartościowe wyniki dotyczące analizy zamożności i dochodowości zarówno polskich gospodarstw domowych jak też gospodarstw domowych społeczeństw należących do Unii Europejskiej traktowanej jako całość. Analizę tę poprowadziłem w oparciu z jednej strony o dane empiryczne (z lat 2004–2010), z drugiej strony o wybrane modele rozwijane w ramach fizyki statystycznej, matematyki finansowej i ekonometrii dotyczące np. funkcji potęgowych, wykładniczych i logarytmiczno–normalnych, czyli (mówiąc ogólniej) w ramach prawa Pareto, modelu zderzeń, Prawa Efektów Proporcjonalnych jak też w ramach uogólnionego modelu Lotka–Volterra. Należy podkreślić, że wspomniane modele mają charakter wielociałowy i dynamiczny w sensie dynamiki stochastycznej, której szczególnym przypadkiem jest liniowa i nieliniowa dynamika Langevina.

Celem merytorycznym mojej rozprawy doktorskiej jest pogłębione zrozumienie, a w tym opis ilościowy (sparametryzowany niewielką liczbą wielkości możliwych do wyznaczenia z danych empirycznych) mechanizmów uzyskiwania dochodów i bogacenia się lub ubożenia zarówno ludności Polski jak i Unii Europejskiej. Jest to niezwykle ważny cel, którego osiągnięcie pociąga za sobą możliwość określenia ścieżek rozwojowych Unii Europejskiej a w tym Polski.

Głównym wynikiem moich badań jest uogólnienie znanego modelu Yakovenko a co za tym idzie zasadniczego wzoru Yakovenko. Umożliwiło to zbadanie dochodowości wszystkich klas społecznych a nie tylko dwóch jak w oryginalnym modelu Yakovenko. Innymi słowy, otrzymana przeze mnie formuła jest w stanie opisać dochodowość wszystkich gospodarstw domowych w Unii Europejskiej, poczynając od najbiedniejszych na najbogatszych kończąc. Należy podkreślić, że pomimo wielu prób mających na celu teoretyczny opis empirycznych rozkładów dochodów jednostek bądź gospodarstw domowych, jak dotąd zaproponowany przeze mnie model jest pierwszym dającym efektywny, spójny i kompleksowy opis dochodów gospodarstw domowych wszystkich klas społecznych. Jest on ukoronowaniem obecnego stanu wiedzy dotyczącej analizy dochodów gospodarstw domowych.

Przedstawione przeze mnie badania naukowe mają ważny aspekt pragmatyczny, gdyż mogą wspomóc np. konstruowanie optymalnego systemu podatkowego, a także emerytalnego. Ponadto, mogą pomóc określić strukturę popytową rynków.

Niniejsza rozprawa doktorska składa się z czterech wzajemnie uzupełniających się rozdziałów:

- w pierwszym rozdziale przedstawiłem krytyczne omówienie stosowanych dotychczas modeli teoretycznych, mających na celu analityczny opis dochodów społeczeństw;
- w drugim rozdziale zaproponowałem własne modele, mające na celu efektywny, spójny i kompleksowy opis dochodów społeczeństw;
- w trzecim rozdziale dokonałem opracowania danych empirycznych dotyczących rocznych dochodów rozporządzalnych gospodarstw domowych w Polsce w latach 2004–2010, przede wszystkim poprzez porównanie ich z modelami teoretycznymi. Skupiłem się także na analizie własności wykorzystywanych modeli;
- w czwartym rozdziale szczegółowo przedstawiłem moją procedurę dotyczącą łączenia różnych baz zawierających dane o dochodach. Następnie, mając pełne dane o dochodach wszystkich klas społecznych w Unii Europejskiej, przeprowadziłem analogiczne analizy jak w rozdziale 3 dla lat 2005–2010. Dokonałem również weryfikacji i obszernego omówienia własności skonstruowanego przeze mnie modelu.

Pracę uzupełniłem o niezbędne rozważania dotyczące rozkładu Pareto (dodatek A i B) oraz uogólnionego modelu Lotka–Volterra (dodatek C). Dla ułatwienia zamieściłem także opis metody Weibulla (dodatek D). W dodatkach E i F zawarłem uzupełniające informacje na temat analizy danych empirycznych dotyczących dochodów gospodarstw domowych w Unii Europejskiej. Z kolei, w dodatku G przeprowadziłem poszerzoną dyskusję na temat procedury łączenia baz danych. Natomiast, dodatek H zawiera omówienie współczynnika Giniego.

ABSTRACT

My phd thesis is an original selection, reformulation and extension of the canonical models of statistical physics dedicated for study the properties of systems beyond the traditional physics. Although, the research areas covered by my thesis are traditionally the domain of economy and sociology, these areas belongs today to the research mainstream of econo- and sociophysics.

More precisely, the subject of my thesis concerns the dynamics of household incomes in Poland and the European Union during the time of crisis and beyond. I conducted my main analysis both within the Langevin stochastic dynamics and Fokker–Planck equation. In addition, I considered selected models of random walks where household income is treated as a time-dependent random variable.

In my thesis I obtained valuable results by analysing the wealth and income of Polish households as well as households of societies belonging to the European Union. This analysis was based both on empirical data (from the years 2004–2010) and on selected models developed within the statistical physics, financial mathematics and econometrics concerning for example, power-law, exponential and log-normal distributions that is, the Pareto law, collision models, the Rules of Proportionate Growth and Generalised Lotka–Volterra model. It should be emphasised that these models are multivariable and dynamic (in the sense of stochastic dynamics) – the linear and nonlinear Langevin dynamics are their prominent examples.

The main aim of my thesis is a deeper understanding, including quantitative description (parametrised by a small number of values possible to determine from the empirical data) of the mechanisms of gaining income and enrichment or impoverishment of the societies both Poland and the European Union. This is an extremely important goal, the achievement of which can specify the pathways of development of the European Union and Poland.

The main result of my research is a generalisation of the well known Yakovenko model and thus the principal formula of Yakovenko. The formula is based directly on the equilibrium solution of the corresponding Fokker–Planck equation. This allowed us to study the income of all three society classes and not only two ones as the original Yakovenko model made. In other words, I have obtained the formula which is able to describe the income of all European Union households belonging to the low-, medium- and high-income society classes. It should be emphasised that, despite many attempts aimed for theoretical description of the empirical income distributions of individuals or households, so far I have proposed model which is the first that gives an effective, consistent and comprehensive description of the household incomes of all society classes. My low-parameter model is a culmination of the current knowledge regarding the analysis of household incomes.

My research has some significant pragmatic aspect, as it can support, for instance, the construction of the appropriate tax and retirement systems. Furthermore, my research can estimate the structure of the market demands.

My dissertation consists of four complementary chapters:

- the first chapter, where I presented a critical overview of theoretical models used so far to analytical description of society incomes;

- the second chapter, where I developed efficient, consistent and comprehensive formalism describing the society incomes;
- the third chapter, where I analysed empirical data on the annual disposable income of Polish households in the years 2004–2010, in particular by comparison of the empirical data with theoretical models. In this chapter I also focused on the analysis of properties of commonly used models;
- the fourth chapter, where I presented the detailed procedure for joining different databases containing empirical income data. Then, with a complete income dataset of all society classes in the European Union, I have carried out an extended (in comparison with that shown in Section 3) analysis for the years 2005–2010. To perform this analysis I used my models, having well motivated opportunity to make their thorough review and discuss their useful properties.

In seven appendices I discussed the Pareto distribution (Appendix A and B) and the Generalised Lotka–Volterra model (Appendix C). I also enclosed a description of the Weibull method (Appendix D). In the appendices E and F I included additional information on the analysis of empirical data on household incomes in the European Union. Furthermore, in the Appendix G I carried out a thorough discussion concerning the procedure of joining the databases. Moreover, Appendix H contains an overview of the Gini coefficient.

SPIS TREŚCI

Wstęp: krótki przegląd problematyki		7
1	MODELE DYNAMIKI DOCHODÓW ZNANE DOTYCHCZAS	13
1.1	Dochód jako zmienna losowa	13
1.2	Przejście od równania Langevina do równania Fokkera–Plancka	14
1.3	Prawo Boltzmann–Gibbsa	24
1.4	Prawa Pareto	30
1.5	Model Yakovenko	33
1.6	Prawo Efektów Proporcjonalnych	35
1.7	Uogólniony model Lotka–Volterra	37
2	MODELE DYNAMIKI DOCHODÓW ROZWINIĘTE W NINIEJSZEJ ROZPRAWIE	43
2.1	Prawo Pareto – model drabinowy	43
2.2	Rozszerzony model Yakovenko	48
3	ANALIZA ROCZNYCH DOCHODÓW GOSPODARSTW DOMOWYCH W POLSCE	53
3.1	Opis bazy danych	53
3.2	Empiryczne dopełnienia dystrybuant	54
3.3	Analiza danych empirycznych	58
3.4	Podsumowanie rozdziału	81
4	ANALIZA ROCZNYCH DOCHODÓW GOSPODARSTW DOMOWYCH W UNII EUROPEJSKIEJ	83
4.1	Opis baz danych	83
4.2	Procedura łączenia baz danych	86
4.3	Empiryczne dopełnienia dystrybuant	89
4.4	Analiza danych empirycznych	93
4.5	Dodatkowe uwagi	116
4.6	Podsumowanie rozdziału	126
Zakończenie		129
A	WŁASNOŚCI ZMIENNYCH LOSOWYCH Z ROZKŁADU PARETO	135
B	ROZKŁAD SUMY N NIEZALEŻNYCH ZMIENNYCH LOSOWYCH POCHODZĄCYCH Z ROZKŁADU PARETO	137
C	ELIMINACJA WSPÓŁCZYNNIKA $c(m_1, \dots, m_N, t)$ W UOGÓLNIONYM MODELU LOTKA–VOLTERRA	141
D	METODA WEIBULLA	145
E	OPIS ZMIENNEJ „ROCZNY DOCHÓD BRUTTO GOSPODARSTWA DOMOWEGO”	147

F	ANALIZA DANYCH EMPIRYCZNYCH – BADANIE EU-SILC	149
G	DODATKOWE UWAGI NA TEMAT PROCEDURY ŁĄCZENIA BAZ DANYCH	151
H	WSPÓŁCZYNNIK GINIEGO	159
	Bibliografia	159

WSTĘP: KRÓTKI PRZEGLĄD PROBLEMATYKI

*„Wiedza nigdy nie rozstrzygnie jednego problemu
bez stworzenia dziesięciu nowych.”*

GEORGE BERNARD SHAW

Ekonofizyka jako dziedzina nauki

Dzisiejsze badania naukowe wykorzystują do opisu złożoności otaczającego nas świata nie tylko narzędzia, jakie oferują poszczególne dyscypliny naukowe, ale coraz częściej łączą metody stosowane do tej pory w zupełnie różnych dziedzinach nauki. Jedną z takich dziedzin nauki jest ekonofizyka, która przy użyciu metod i modeli stosowanych najczęściej w fizyce statystycznej i fizyce materii skondensowanej opisuje procesy ekonomiczne i finansowe. Terminu *ekonofizyka* po raz pierwszy użył fizyk Eugene H. Stanley w pracy [1], która stanowi swego rodzaju manifest tej nowej dziedziny nauki.

Termin *ekonofizyka* powstał jako neologizm – złożenie dwóch słów: ekonomia i fizyka określających dwie odrębne dziedziny nauki, analogicznie jak ma to miejsce dla astrofizyki, geofizyki czy biofizyki, które wykorzystują metody fizyki do opisu zjawisk badanych odpowiednio w ramach astronomii, nauki o Ziemi czy biologii. Należy podkreślić, że ekonofizyka nie przenosi bezpośrednio praw fizyki, np. takich jak zasady dynamiki Newtona czy podejścia stosowanego w ramach fizyki kwantowej, do opisu zachowań ekonomicznych różnego rodzaju podmiotów, np. inwestorów, poszczególnych osób czy gospodarstw domowych. Wykorzystuje ona metody matematyczne rozwinięte np. w ramach fizyki statystycznej do analizy własności statystycznych układów złożonych składających się z dużej liczby wspomnianych wcześniej podmiotów [2].

Ekonofizyka skupia się na ilościowej analizie danych ekonomicznych i finansowych, które w dobie komputeryzacji i internetu stały się łatwo dostępne. Koncen-

trując się na modelowaniu matematyczno-fizycznym dużej liczby oddziałujących ze sobą ekonomicznych podmiotów (ang. *agents*), ma ona wiele wspólnego z badaniami prowadzonymi w ramach ekonometrii i modelowania wieloagentowego (ang. *agent-based modeling*) [2].

Aktualny stan wiedzy

Jednym z ważniejszych nurtów ekonofizyki jest badanie redystrybucji bogactwa i dochodów w społeczeństwie oraz analiza nierówności społecznych. Pionierem tego typu badań jest włoski ekonomista i socjolog – Vilfredo Pareto. Pod koniec XIX wieku Pareto jako pierwszy podjął próbę analitycznego opisu dystrybucji bogactwa w społeczeństwie reprezentowanego przez roczne dochody jednostek. Jednym z ważniejszych jego odkryć jest fakt, że rozkłady dochodów ludności poszczególnych państw mają charakter uniwersalny i cechują się małą zmiennością w „przestrzeni” i czasie. Dodatkowo stwierdził on, że rozkłady te nie przypominają kształtem rozkładów, które otrzymalibyśmy gdyby zdobywanie i posiadanie bogactwa miało charakter przypadkowy. Pareto zaobserwował także stabilność tych rozkładów. Nawet gdybyśmy wykluczyli z procesu uzyskiwania dochodów jednostki najbogatsze bądź najbiedniejsze, to po pewnym okresie rozkład dochodów będzie ponownie przypominał kształtem rozkład początkowy [3–6]. Głównym wynikiem jego badań było stwierdzenie, że rozkłady dochodów w poszczególnych krajach o stabilnej gospodarce mogą być opisane uniwersalnym prawem potęgowym, zwanym dzisiaj prawem Pareto (szczegółowy opis prawa Pareto znajduje się w rozdziale 1.4). Jako możliwe pochodzenie tego prawa, Pareto wskazał samopodobieństwo struktury dochodów uzyskiwanych przez społeczeństwa. Odkrycia Pareto sprawiły, że badanie i próba analitycznego opisu dochodów społeczeństw wywołały istną lawinę prac naukowych.

Tematyką dochodów społeczeństw zajmował się także francuski ekonomista Robert Gibrat. W swoich badaniach wskazał on, że prawo Pareto nie jest w stanie opisać rozkładu dochodów w całym zakresie. Próbuąc stworzyć model, który opisywałby dochody wszystkich klas społecznych, zaproponował podejście zwane Prawem Efektów Proporcjonalnych. Wykorzystując proces stochastyczny opisujący dynamikę zmian dochodu lub bogactwa pojedynczej osoby bądź gospodarstwa

domowego, otrzymał teoretyczną gęstość prawdopodobieństwa opisującą dochody społeczeństwa (patrz rozdział 1.6) [5–10].

Z kolei, David Champernowne zaproponował jeden z pierwszych modeli stochastycznych, który odtwarza prawo Pareto [6, 11]. Natomiast Benoit Mandelbrot opisał podstawowe własności zmiennych losowych pochodzących z rozkładu Pareto (patrz dodatek A) [6, 12].

Badanie i próba analitycznego opisu rozkładów dochodów ludności podjęta przez Gibrata, Champernowne’a oraz Mandelbrota doprowadziła do odkrycia wielu istotnych własności tych rozkładów, jednak nie dała odpowiedzi na pytanie o mechanizm „mikroskopowy” determinujący określoną postać rozkładu dochodów. Współcześnie zaproponowano modele, przy pomocy których próbuje się wyjaśnić mechanizmy mikroskopowe (na poziomie dochodów poszczególnych osób bądź gospodarstw domowych) stojące za obserwowanymi rozkładami dochodów.

Wśród tych modeli możemy wyróżnić dwie zasadnicze grupy. Pierwsza opiera się na teorii kinetyczno-cząsteczkowej Boltzmanna opisującej zderzenia cząsteczek w gazach. Analogicznie do zderzenia dwóch cząsteczek, podczas którego następuje wymiana energii kinetycznej pomiędzy nimi, w modelach tych – zwanych modelami zderzeń – zakłada się, że w danej chwili dwie losowo wybrane osoby lub gospodarstwa domowe dokonują pomiędzy sobą wymiany posiadanych zasobów finansowych zgodnie z określonymi regułami [13–21]. Wybierając odpowiednio reguły możemy mówić o podstawowym modelu zderzeń zaproponowanym przez Drăgulescu i Yakovenko [6, 19, 22], który prowadzi do prawa Boltzmanna–Gibbsa (patrz rozdział 1.3). Wyróżnia się także modele zderzeń dopuszczające ujemny dochód (rozumiany jako strata) [19, 23–25], modele zderzeń z jednakową dla wszystkich jednostek skłonnością do oszczędzania [6, 13–17, 20, 22, 26] a także modele zderzeń ze zróżnicowaną skłonnością jednostek do oszczędzania [6, 18, 20–22, 27–36].

Druga grupa modeli, rozwijanych również w ramach niniejszej rozprawy doktorskiej, traktuje dochód jednostki bądź gospodarstwa domowego jako zmienną losową, do opisu której stosuje się dynamikę stochastyczną wyrażoną przez nieliniowe równanie Langevina i odpowiadające mu równanie Fokkera–Plancka. W zależności od konkretnych założeń dotyczących dynamiki zmian dochodu, możemy otrzymać następujące modele: prawo Boltzmanna–Gibbsa (patrz rozdział 1.3) [2,

5, 6, 37], prawo Pareto (rozdział 1.4) [2, 5, 6, 37], Prawo Efektów Proporcjonalnych (rozdział 1.6) [2, 5–10], uogólniony model Lotka–Volterra (rozdział 1.7) [2, 5, 38–41], bądź model Yakovenko (rozdział 1.5) [2, 37].

Należy podkreślić, że zarówno w ramach teorii kinetyczno-cząsteczkowej Boltzmanna jak i dynamiki stochastycznej Langevina, które prowadzą do określonych rozkładów dochodów jednostek bądź gospodarstw domowych, ekonofizyka toruje drogę nowym kierunkom badań, komplementarnym do rozwijanych w ramach nauk ekonomicznych [42, 43].

Oprócz budowy modeli analitycznych opisujących rozkłady dochodów, wielu badaczy zajmowało się również weryfikacją tych modeli poprzez porównanie ich przewidywań z danymi empirycznymi dotyczącymi zarówno dochodów jednostek, jak też dochodów gospodarstw domowych. Analizy takie przeprowadzono m. in. dla Stanów Zjednoczonych [2, 30, 37, 44–55], Wielkiej Brytanii [3, 5, 29–31, 45, 52, 56, 57], Niemiec [3, 52, 57], Włoch [3, 57–59], Francji [3], Szwajcarii [3], Japonii [29, 53, 56, 60–64], Australii [54, 59, 65, 66], Kanady [47], Czech [47], Nowej Zelandii [29, 56], Indii [67], Sri Lanki [47], Argentyny [56], Peru [3], Korei Południowej [68], Rumunii [69] oraz Polski [49].

Należy zaznaczyć, że (przynajmniej częściowo) tematyka niniejszej pracy dotyczy także zagadnień rozwijanych w ramach socjofizyki. Socjofizyka w odróżnieniu od ekonofizyki nie skupia się tylko i wyłącznie na badaniach ekonomicznej aktywności jednostek, ale wykorzystując metody fizyki, zajmuje się szeroko rozumianymi zagadnieniami z dziedziny nauk społecznych, takimi jak np. analiza preferencji wyborczych, sieci społecznych, decyzji podejmowanych w ramach określonych grup społecznych, formowanie koalicji, terroryzm czy dynamika opinii społecznych [2, 70].

Cele pracy

Niniejsza rozprawa doktorska opiera się głównie na moich pięciu publikacjach [71–75], z lat 2010–2013, rozszerzając znacząco oraz uzupełniając i pogłębiając badania zapoczątkowane przeze mnie w pracy magisterskiej [6].

Głównym celem niniejszej pracy jest:

- skonstruowanie modelu teoretycznego (nazwanego przeze mnie rozszerzonym modelem Yakovenko, rozdział 2.2) opisującego rozkłady dochodów w całym zakresie, czyli dochody wszystkich klas społecznych, tj. gospodarstw domowych o niskich, średnich i wysokich dochodach, włączając w to także obszary przejściowe pomiędzy tymi klasami. Należy podkreślić, że pomimo wspomnianych wcześniej wielu prób teoretycznego opisu empirycznych rozkładów dochodów jednostek bądź gospodarstw domowych jest to, jak dotąd, pierwszy model dający efektywny, spójny i kompleksowy opis dochodów gospodarstw domowych wszystkich klas społecznych. Jest on ukoronowaniem obecnego stanu wiedzy dotyczącej analizy dochodów gospodarstw domowych.

Wśród celów szczegółowych należy wymienić:

- wyprowadzenie równania Fokkera–Plancka z równania Langevina dostosowane do konkretnych warunków brzegowych postawionego w niniejszej pracy problemu;
- zaproponowanie procedury łączenia baz danych, dzięki której możliwe było uzyskanie wystarczającej liczby danych o dochodach należących do wszystkich klas społecznych Unii Europejskiej, tj. gospodarstw domowych uzyskujących niskie, średnie i wysokie dochody (rozdział 4.2);
- analizę danych empirycznych dotyczących rocznych dochodów rozporządzalnych gospodarstw domowych w Polsce polegającą na weryfikacji przytoczonych w pracy modeli: Prawa Efektów Proporcjonalnych, słabego prawa Pareto oraz uogólnionego modelu Lotka–Volterra (rozdział 3.3);
- analizę danych empirycznych dotyczących rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej polegającą na weryfikacji przytoczonych w pracy modeli: prawa Boltzmann–Gibbsa, słabego prawa Pareto a przede wszystkim weryfikacji skonstruowanego przeze mnie rozszerzonego modelu Yakovenko (rozdział 4.4 oraz 4.5). Ku mojemu zaskoczeniu,

analiza dochodów gospodarstw domowych Unii Europejskiej metodami fizyki statystycznej została po raz pierwszy przeprowadzona w moich publikacjach [73–75] i niniejszej rozprawie doktorskiej.

Układ pracy

Niniejsza praca składa się z czterech wzajemnie uzupełniających się rozdziałów:

- w pierwszym rozdziale omówiłem stosowane dotychczas modele teoretyczne, mające na celu analityczny opis dochodów społeczeństw;
- w drugim rozdziale zaproponowałem własne modele, mające na celu efektywny, spójny i kompleksowy opis dochodów społeczeństw;
- w trzecim rozdziale dokonałem opracowania danych empirycznych dotyczących rocznych dochodów rozporządzalnych gospodarstw domowych w Polsce w latach 2004–2010, przede wszystkim poprzez porównanie ich z modelami teoretycznymi. Skupiłem się także na analizie własności wykorzystywanych modeli;
- w czwartym rozdziale szczegółowo przedstawiłem moją procedurę dotyczącą łączenia różnych baz zawierających dane o dochodach. Następnie, mając pełne dane o dochodach wszystkich klas społecznych w Unii Europejskiej, wykonałem analogiczne analizy jak w rozdziale 3 dla lat 2005–2010. Dokonałem również weryfikacji i obszernego omówienia własności skonstruowanego przeze mnie modelu.

Pracę uzupełniłem o niezbędne rozważania dotyczące rozkładu Pareto (dodatek A i B) oraz uogólnionego modelu Lotka–Volterra (dodatek C). Zamieściłem także opis metody Weibulla (dodatek D). W dodatkach E i F zawarłem uzupełniające informacje na temat analizy danych empirycznych dotyczących dochodów gospodarstw domowych w Unii Europejskiej. Z kolei, w dodatku G przeprowadziłem poszerzoną dyskusję na temat procedury łączenia baz danych. Natomiast, dodatek H zawiera omówienie współczynnika Giniego.

MODELE DYNAMIKI DOCHODÓW ZNANE DOTYCHCZAS

*„This opens a broad perspective,
if we do not only think of mechanical objects.
Let's consider to apply this method
to the statistics of living beings,
society, sociology and so forth.”*

LUDWIG EDUARD BOLTZMANN

W rozdziale tym omawiam pojęcia, twierdzenia oraz modele znane i stosowane dotychczas do opisu dochodów gospodarstw domowych. Dopiero w kolejnym rozdziale 2 przedstawię modele rozwinięte w ramach niniejszej rozprawy doktorskiej.

1.1 DOCHÓD JAKO ZMIENNA LOSOWA

Dochód gospodarstwa domowego jest zmienną, która w ogólności zależy od czasu w sposób niezwykle skomplikowany co powoduje, że charakterystyka jego dynamiki nie została, jak dotąd, przedstawiona w sposób zadawalający. Komplikacja ta wynika z faktu, że obok zależności o charakterze deterministycznym istotną rolę w jego dynamice odgrywa czynnik losowy, a zatem na wielkość dochodu ma istotny wpływ niepewność. Innymi słowy, dochód można traktować jak zmienną losową a do opisu jego dynamiki stosować dynamikę stochastyczną. W rezultacie, ewolucją dochodów możemy zajmować się w ramach teorii procesów stochastycznych. Należy podkreślić, że niedeterministyczna składowa dochodów bierze się

stąd, że gospodarstwo domowe może mieć inne (niż etatowe, zakontraktowane), nieprzewidywalne dochody np. nagle pojawiająca się praca zlecona, prace dorywcze, czy zysk z inwestycji giełdowych. To znaczy, pracownicy etatowi mają okazję do osiągnięcia dodatkowych, przypadkowych dochodów.

Stosując podejście fizyki statystycznej możemy, zamiast określać jak zmienia się dochód poszczególnych gospodarstw domowych (np. w danym kraju, bądź grupie społecznej), mówić o prawdopodobieństwie tego, że dane gospodarstwo domowe uzyska dochód w zadanym przedziale wartości.

Strategia niniejszej pracy polega na tym, aby wychodząc od nieliniowego równania Langevina opisującego konkretną dynamikę stochastyczną, tzn. dynamikę dochodów gospodarstw domowych, dojść do równania Fokkera–Plancka, a stąd uzyskać jego równowagowe rozwiązanie pozwalające opisać roczne dochody gospodarstw domowych. Na tej drodze zamierzam osiągnąć zasadniczy cel pracy jakim jest zbudowanie modelu opisującego dochody gospodarstw domowych należących do wszystkich grup społecznych, poczynając od najbiedniejszych na najbogatszych kończąc.

W niniejszym rozdziale przedstawiam najważniejsze modele dynamiki dochodów, aby na tym tle jasno uwypuklić cele i osiągnięcia pracy.

1.2 PRZEJŚCIE OD RÓWNANIA LANGEVINA DO RÓWNANIA FOKKERA–PLANCKA

Jak wiadomo, przejście od równania Langevina do równania Fokkera–Plancka przerzuca pomost pomiędzy dynamiką stochastyczną a procesami stochastycznymi – stanowi to strategiczny element niniejszej pracy. Przedstawiam je tutaj aby podkreślić, że przejście to jest na tyle ogólne, że stosuje się zarówno do zmiennej losowej posiadającej konkretną interpretację (tutaj dochodu gospodarstwa domowego), jak też do konkretnej postaci współczynników $A(m)$ oraz $C(m)$ pojawiających się w równaniu Langevina. Co więcej, **wyprowadzenie tutaj równania Fokkera–Plancka z równania Langevina wymagało dostosowania tego wyprowadzenia do konkretnych warunków brzegowych narzuconych przez postawiony w niniejszej pracy problem.**

Rozważmy zależny od czasu dochód $m(t)$ gospodarstwa domowego, który w czasie Δt może ulec zmianie. Jak zostało wcześniej powiedziane, dochód ten można traktować jako zmienną losową. Zatem, ewolucję czasową tej zmiennej możemy opisać przy pomocy dyskretnego nieliniowego równania Langevina [76, 77]:

$$\frac{m(t + \Delta t) - m(t)}{\Delta t} = -A(m(t)) + C(m(t)) \eta(t), \quad (1.1)$$

gdzie przyjmujemy, że Δt jest rzędu jednego dnia (dochód uzyskiwany w postaci tzw. dniówki), natomiast czas t jest rzędu jednego miesiąca. Ponieważ $\Delta t \ll t$, więc możemy wykonać przejście graniczne $\Delta t \rightarrow 0$ i zapisać dyskretne równanie Langevina w postaci uciągłonej:

$$\frac{dm(t)}{dt} = -A(m(t)) + C(m(t)) \eta(t). \quad (1.2)$$

W powyższym równaniu $A(m(t))$ jest (w ogólności) nieliniowym współczynnikiem dryfu a $\eta(t)$ jest zależnym od czasu białym szumem¹; amplituda tego szumu $C(m(t))$ może być także nieliniową funkcją $m(t)$. Inaczej mówiąc, jest to równanie Langevina z siłą oporu lepkiego równą $-A(m(t))$ oraz siłą stochastyczną równą $C(m(t)) \eta(t)$. Przy czym, dla siły stochastycznej spełnione są następujące warunki:

$$\langle \eta(t) \rangle = 0, \quad (1.3)$$

$$\langle \eta(t) \eta(t') \rangle = \delta(t - t'). \quad (1.4)$$

Z punktu widzenia ekonomii, człon $-A(m(t))$ wyraża deterministyczną część dochodów uzyskiwanych przez gospodarstwo domowe (np. pensje, renty, emerytury), a człon $C(m(t)) \eta(t)$ wyraża niedeterministyczną, nieprzewidywalną część dochodów pochodzących z pracy dorywczej, bądź stanowiących np. zyski z inwestycji giełdowych.

Pojawiający się w równaniu (1.2) minus przed współczynnikiem $A(m(t))$ wymaga dodatkowego komentarza. Minus ten sprawia, że równoważne mu równanie Fokkera–Plancka posiada rozwiązania równowagowe. Oczywiście, gdyby pominąć w tym równaniu składnik stochastyczny $C(m(t)) \eta(t)$ to wówczas, mielibyśmy do czynienia z sytuacją nieustannego malenia dochodów nominalnych (które

¹ Możliwe jest również uogólnienie przedstawionego podejścia polegające na tym, że $\eta(t)$ jest opisywane dowolnym rozkładem stabilnym [78, 79].

to dochody będą analizowane w rozdziale 3 oraz 4). Byłaby to sytuacja nierealna. Zatem, ten składnik musi być uwzględniony. Równanie (1.2) należy zawsze rozpatrywać w kontekście występowania dwóch składników: deterministycznego $-A(m(t))$ i stochastycznego $C(m(t))\eta(t)$. Pozwala to skonstruować rozwiązanie równowagowe.

Analizując dochody gospodarstw domowych przy użyciu równania (1.2) będziemy poszukiwać rozwiązań długookresowych (równowagowych), przy czym dla przyjętych skal czasowych zakładamy, że dynamika zmian dochodu osiąga stacjonarność w okresie jednego roku. Zatem minus przed $A(m(t))$ może wynikać z faktu, że nominalne dochody maleją. Malenie nominalnych dochodów (w przedziale jednego roku) jest możliwe w czasach kryzysu z niską inflacją. Wówczas mamy do czynienia z sytuacją cięć zarówno w sektorze państwowym (np. obniżenie stałych zasiłków z pomocy społecznej) jak i w sektorze prywatnym (np. cięcia pensji). Założenia te wydają się sensowne w obecnej sytuacji gospodarczej Polski i Unii Europejskiej.

Dodatkowo należy zauważyć, że dochody gospodarstw domowych nie mogą rosnać do nieskończoności (gdyby tak było to nie moglibyśmy w ogóle poszukiwać stacjonarności). Zatem „minus” wyraża fakt, że dochody gospodarstw domowych nie rosną w nieskończoność.

Poniżej zostanie przedstawiona metoda przejścia od równania Langevina (1.2) do równania Fokkera–Plancka, która została zainspirowana książką² [77]. Jednakże, podejście zaproponowane w niniejszej pracy różni się istotnie od dotychczasowych przyjętymi warunkami brzegowymi, o czym jest mowa poniżej.

Przejście od równania Langevina do równania Fokkera–Plancka

Wyprowadzenie równania Fokkera–Plancka z równania Langevina składa się z kilku etapów. W pierwszym etapie odcałkowujemy równanie dynamiki stocha-

² Jest to nadzwyczaj ważne wyprowadzenie, którego brak w książce [76] uważanej za kultową dla tej tematyki. Niniejsze wyprowadzenie dotyczy sytuacji nieco ogólniejszych od tych przedstawianych dotychczas w literaturze przedmiotu ([77] oraz referencje tamże).

stycznej Langevina rządzone przez biały szum η , tak aby uzyskać równanie dynamiki stochastycznej rządzone przez szum Wienera (gaussowski). Mianowicie³:

$$dm = -A(m)dt + C(m)dW, \quad (1.5)$$

gdzie $dW = \int_t^{t+dt} \eta(t')dt'$ jest procesem Wienera, co oznacza, że typowe wartości $(dW)^2$ są rzędu dt . Innymi słowy, dW jest odlosowywane z rozkładu Gaussa o wariancji proporcjonalnej do dt . Dodajmy, że w powyższym równaniu przyjęliśmy $A(m)$ i $C(m)$ z początku przedziału czasu $[t, t + dt]$. Tym samym, wyprowadzenie zostanie przedstawione w reprezentacji Itô [76, 77].

W następnym etapie korzystamy z lematu Itô, który sprowadza się do następującego rozwinięcia dowolnej, różniczkowalnej co najmniej dwukrotnie funkcji $h(m)$:

$$dh = -\frac{dh}{dm}A(m)dt + \frac{d^2h}{dm^2}\frac{C^2(m)}{2}dt + \frac{dh}{dm}C(m)dW. \quad (1.6)$$

Przy wyprowadzeniu lematu Itô wykorzystaliśmy bezpośrednio równanie dynamiki stochastycznej (1.5)⁴.

W kolejnym etapie średniujemy równanie (1.6) względem procesu Wienera dW uzyskując:

$$\begin{aligned} \frac{d\langle h \rangle}{dt} &= \left\langle -\frac{dh}{dm}A(m) \right\rangle + \left\langle \frac{d^2h}{dm^2}B(m) \right\rangle = \\ &= \int_{m_0}^{m_{\max}} \left[-\frac{dh}{dm}A(m) + \frac{d^2h}{dm^2}B(m) \right] P(m, t) dm, \end{aligned} \quad (1.7)$$

gdzie, w związku z tym, wyraz zawierający dW wyzerował się; ponadto, wykorzystaliśmy równość $\langle dh \rangle = d\langle h \rangle$ oraz przyjęliśmy, że $B(m) = \frac{C^2(m)}{2}$. W literaturze przedmiotu (dotyczącej równań stochastycznych) przyjmuje się, że $m_0 = -\infty$, a $m_{\max} = \infty$; tutaj tak nie jest, gdyż nie rozważamy strat, ani nieskończonych wartości dochodu. Zatem, należy przyjąć, że minimalna wartość dochodu $m_0 = 0$, a jego maksymalna wartość $m_0 < m_{\max} < \infty$, innymi słowy, całkujemy po ograniczonym nośniku.

³ W dalszych rozważaniach, w celu uproszczenia zapisu przyjmujemy $m(t) \equiv m$.

⁴ W przeprowadzonym rachunku wykorzystaliśmy typową wartość $(dW)^2 = dt$, oraz następującą postać różniczki zupełnej: $dh = \frac{dh}{dm} dm + \frac{1}{2} \frac{d^2h}{dm^2} (dm)^2$.

Całkując przez części prawą stronę powyższego równania i zakładając znikanie funkcji $h(m)$ oraz jej pochodnej dla brzegowych wartości dochodów uzyskiwanych przez gospodarstwa domowe, otrzymujemy:

$$\frac{d \langle h \rangle}{dt} = \int_{m_0}^{m_{\max}} h(m) \left\{ \frac{\partial}{\partial m} [A(m)P(m, t)] + \frac{\partial^2}{\partial m^2} [B(m)P(m, t)] \right\} dm. \quad (1.8)$$

Z drugiej strony, pochodna po czasie wartości średniej wynosi:

$$\frac{d \langle h \rangle}{dt} = \frac{d}{dt} \int_{m_0}^{m_{\max}} h(m)P(m, t) dm = \int_{m_0}^{m_{\max}} h(m) \frac{\partial}{\partial t} P(m, t) dm. \quad (1.9)$$

Porównując lewe strony równań (1.8) i (1.9) oraz pamiętając, że są one sobie równe dla dowolnej funkcji $h(m)$, otrzymujemy ostatecznie poszukiwane równanie⁵:

$$\frac{\partial P(m, t)}{\partial t} = \frac{\partial}{\partial m} [A(m)P(m, t)] + \frac{\partial^2}{\partial m^2} [B(m)P(m, t)]. \quad (1.10)$$

Jest to równanie Fokkera–Plancka (czasem zwane ogólnym równaniem dyfuzji). W ogólności zbiór możliwych wartości m powinien być ciągły, a współczynniki $A(m)$ oraz $B(m)$ mogą być dowolnymi różniczkowalnymi funkcjami rzeczywistymi, przy czym powinno zachodzić $B(m) > 0$, tak aby rozwiązanie równania (1.10) było poprawnie zdefiniowanym rozkładem prawdopodobieństwa [76]. Należy podkreślić, że powyższe wyprowadzenie zostało dostosowane do warunków narzuconych przez postawiony w niniejszej pracy problem – ma specyficzny charakter. Dlatego właśnie zostało przedstawione w całości.

Przejsie od procesu stochastycznego Markowa do równania Fokkera–Plancka

Równanie (1.2) – w rozumieniu fizyki statystycznej – wyraża błądzenie przypadkowe (losowe) w ośrodku lepkim, które na poziomie procesu stochastycznego można opisać przy pomocy równania łańcuchowego Bachelier’a–Chapmana–Kołmogorowa. Porzucamy tutaj na chwilę opis na poziomie dynamiki stochastycznej na rzecz opisu na poziomie procesu stochastycznego Markowa [76, 80–82]. Chodzi nam o jasne pokazanie, że rozważany proces jest pozbawiony pamięci.

⁵ Wspomniane powyżej warunki brzegowe nałożone na funkcję $h(m)$ nie zmniejszają ogólności równania (1.10).

Zatem:

$$\begin{aligned}
 P(m, t + \Delta t | m_0, t_0) &= \sum_{\Delta m} W(m, t + \Delta t | m - \Delta m, t) P(m - \Delta m, t | m_0, t_0) = \\
 &= \sum_{\Delta m \neq 0} W(m, t + \Delta t | m - \Delta m, t) P(m - \Delta m, t | m_0, t_0) + \\
 &+ W(m, t + \Delta t | m, t) P(m, t | m_0, t_0), \tag{1.11}
 \end{aligned}$$

gdzie wielkość $P(m, t | m_0, t_0)$ oznacza prawdopodobieństwo warunkowe (lub gęstość prawdopodobieństwa warunkowego w przypadku ciągłych zmiennych losowych) uzyskania przez gospodarstwo domowe dochodu m w chwili t pod warunkiem, że w chwili początkowej t_0 dochód ten wynosił m_0 . Wielkość $P(m, t | m_0, t_0)$ jest poszukiwaną, kluczową wielkością procesu stochastycznego.

W równaniu (1.11) występuje prawdopodobieństwo przejścia $W(m, t + \Delta t | m - \Delta m, t)$, które dla $\Delta m \neq 0$ oznacza prawdopodobieństwo warunkowe uzyskania przez gospodarstwo domowe dochodu m w chwili $t + \Delta t$ pod warunkiem, że w chwili t uzyskało ono dochód $m - \Delta m$ (gdzie Δm może być tutaj zarówno dodatnie jak i ujemne, a także może znikać). Z kolei, gdy $\Delta m = 0$ prawdopodobieństwo $W(m, t + \Delta t | m, t)$ (tzw. prawdopodobieństwo przetrwania) wyznacza prawdopodobieństwo tego, iż dochód gospodarstwa domowego nie ulegnie zmianie od chwili t do chwili $t + \Delta t$. Prawdopodobieństwa $\{W\}$ jednoznacznie definiują proces stochastyczny Markowa. Spełniają one warunek normalizacyjny postaci:

$$\sum_{\Delta m \neq 0} W(m - \Delta m, t + \Delta t | m, t) + W(m, t + \Delta t | m, t) = 1. \tag{1.12}$$

Z powyższego równania wyznaczamy prawdopodobieństwo przetrwania $W(m, t + \Delta t | m, t)$:

$$W(m, t + \Delta t | m, t) = 1 - \sum_{\Delta m \neq 0} W(m - \Delta m, t + \Delta t | m, t) \tag{1.13}$$

i wstawiamy go do wyrażenia (1.11):

$$\begin{aligned}
 P(m, t + \Delta t | m_0, t_0) &= \sum_{\Delta m \neq 0} W(m, t + \Delta t | m - \Delta m, t) P(m - \Delta m, t) + \\
 &+ (1 - \sum_{\Delta m \neq 0} W(m - \Delta m, t + \Delta t | m, t)) P(m, t | m_0, t_0). \tag{1.14}
 \end{aligned}$$

Wykonując proste przekształcenia algebraiczne i dzieląc stronami przez Δt otrzymujemy:

$$\begin{aligned} & \frac{P(m, t + \Delta t | m_0, t_0) - P(m, t | m_0, t_0)}{\Delta t} = \\ & = \sum_{\Delta m \neq 0} \frac{W(m, t + \Delta t | m - \Delta m, t)}{\Delta t} P(m - \Delta m, t | m_0, t_0) + \\ & - \sum_{\Delta m \neq 0} \frac{W(m - \Delta m, t + \Delta t | m, t)}{\Delta t} P(m, t | m_0, t_0). \end{aligned} \quad (1.15)$$

Korzystając z jednorodności czasu dla prawdopodobieństw warunkowych przyjmujemy, że prawdopodobieństwo przejścia W nie zależy od wartości obu chwil czasu $t + \Delta t$ oraz t , a jedynie od ich różnicy Δt . Zatem, zakładamy że opisywany proces stochastyczny jest stacjonarnym procesem Markowa. Prawdopodobieństwa przejścia $\{W\}$ możemy wówczas przedstawić w zależności od wielkości skoku Δm oraz punktów startowych (tutaj, odpowiednio, $m - \Delta m$ oraz m):

$$W(m, t + \Delta t | m - \Delta m, t) = W(\Delta m, \Delta t | m - \Delta m), \quad (1.16)$$

$$W(m - \Delta m, t + \Delta t | m, t) = W(-\Delta m, \Delta t | m). \quad (1.17)$$

Prawdopodobieństwo $W(\Delta m, \Delta t | m - \Delta m)$ oznacza prawdopodobieństwo zmiany dochodu gospodarstwa domowego o Δm w przedziale czasu Δt pod warunkiem, że wcześniej dochód ten wynosił $m - \Delta m$. Analogicznie interpretujemy prawdopodobieństwo $W(-\Delta m, \Delta t | m)$. Uwzględniając (1.16) i (1.17), równanie (1.15) przyjmuje postać:

$$\begin{aligned} & \frac{P(m, t + \Delta t | m_0, t_0) - P(m, t | m_0, t_0)}{\Delta t} = \\ & = \sum_{\Delta m \neq 0} \frac{W(\Delta m, \Delta t | m - \Delta m)}{\Delta t} P(m - \Delta m, t | m_0, t_0) + \\ & - \sum_{\Delta m \neq 0} \frac{W(-\Delta m, \Delta t | m)}{\Delta t} P(m, t | m_0, t_0). \end{aligned} \quad (1.18)$$

Wykonując obustronnie przejście graniczne $\Delta t \rightarrow 0$, otrzymujemy różniczkowo-różnicowe równanie Mistrza opisujące ewolucję poszukiwanego prawdopodobieństwa $P(m, t | m_0, t_0)$:

$$\begin{aligned} \frac{\partial P(m, t | m_0, t_0)}{\partial t} & = \sum_{\Delta m \neq 0} \Gamma(\Delta m | m - \Delta m) P(m - \Delta m, t | m_0, t_0) + \\ & - \sum_{\Delta m \neq 0} \Gamma(-\Delta m | m) P(m, t | m_0, t_0), \end{aligned} \quad (1.19)$$

gdzie elementy przejścia:

$$\Gamma(\Delta m|m - \Delta m) \stackrel{\text{def.}}{=} \lim_{\Delta t \rightarrow 0} \frac{W(m, t + \Delta t | m - \Delta m, t)}{\Delta t} = \lim_{\Delta t \rightarrow 0} \frac{W(\Delta m, \Delta t | m - \Delta m)}{\Delta t}, \quad (1.20)$$

$$\Gamma(-\Delta m|m) \stackrel{\text{def.}}{=} \lim_{\Delta t \rightarrow 0} \frac{W(m - \Delta m, t + \Delta t | m, t)}{\Delta t} = \lim_{\Delta t \rightarrow 0} \frac{W(-\Delta m, \Delta t | m)}{\Delta t}, \quad (1.21)$$

oznaczają tutaj prawdopodobieństwa przejścia na jednostkę czasu Δt zwane funkcjami intensywności lub po prostu szybkościami procesu stochastycznego. Zostały one uzyskane przy naturalnym założeniu ich liniowej zależności od Δt dla małych Δt . Aby można było rozwiązać równanie (1.19), intensywności procesu muszą być zadane [6, 76, 82], gdyż to one właśnie „rządzą” procesem stochastycznym.

Jeśli dopuścimy możliwość ciągłych zmian wielkości dochodu gospodarstw domowych, to wówczas równanie Mistrza możemy zapisać w postaci:

$$\frac{\partial P(m, t)}{\partial t} = \int \Gamma(s|m - s)P(m - s, t)ds - \int \Gamma(-s|m)P(m, t)ds, \quad (1.22)$$

gdzie $P(m, t|m_0, t_0) \equiv P(m, t)$ a w celu uproszczenia notacji przyjęto, że $s \equiv \Delta m$. Ponieważ $P(m, t)$ nie zależy od s otrzymujemy:

$$\frac{\partial P(m, t)}{\partial t} = \int \Gamma(s|m - s)P(m - s, t)ds - P(m, t) \int \Gamma(-s|m)ds. \quad (1.23)$$

W ogólności, prawdopodobieństwo dużej zmiany dochodu danego gospodarstwa domowego na jednostkę czasu jest stosunkowo niskie, stąd zakładamy że:

$$\Gamma(s|m - s) \approx 0 \quad \text{dla} \quad |s| > s_{\text{prog}}, \quad (1.24)$$

gdzie $s_{\text{prog}} (> 0)$ jest zadaną wielkością progową. Zatem, przyjmujemy że w skali miesięcznej nie następują spektakularne wzrosty, ani spadki dochodów gospodarstw domowych.

Dalej interesuje nas rozwiązanie $P(m, t)$, które zmienia się powoli ze zmianą m , stąd przyjmujemy że:

$$\Gamma(s|m - s + \delta m) \approx \Gamma(s|m - s) \quad \text{dla} \quad |\delta m| < \min(m, s_{\text{prog}}). \quad (1.25)$$

Po uwzględnieniu założeń (1.24) i (1.25), element $\Gamma(s|m - s)P(m - s, t)$ stojący w równaniu (1.23) możemy rozwinąć w szereg Taylora w punkcie m pozostawiając wyrazy do drugiego rzędu włącznie. Otrzymujemy wówczas:

$$\begin{aligned} \frac{\partial P(m, t)}{\partial t} = & P(m, t) \int \Gamma(s|m)ds - \int s \frac{\partial}{\partial m} [\Gamma(s|m)P(m, t)] ds + \\ & + \frac{1}{2} \int s^2 \frac{\partial^2}{\partial m^2} [\Gamma(s|m)P(m, t)] ds - P(m, t) \int \Gamma(-s|m)ds. \end{aligned} \quad (1.26)$$

Oczywiście, można uwzględnić wszystkie wyrazy rozwinięcia w szereg Taylora, ale wówczas otrzymalibyśmy rozwinięcie Kramersa–Moyala, które formalnie jest równoważne równaniu Mistrza (1.22). Zatem, posługiwanie się nim nie jest prostsze a co za tym idzie, utrudnione jest uzyskanie rozwiązania równowagowego w zamkniętej postaci, jaką uzyskuje się z równania Fokkera–Plancka – a to jest przecież celem niniejszego wywodu.

Przyjmując oznaczenia⁶:

$$A(m) = - \int_{-\infty}^{\infty} s \Gamma(s|m) ds \quad (1.27)$$

oraz:

$$B(m) = \frac{1}{2} \int_{-\infty}^{\infty} s^2 \Gamma(s|m) ds, \quad (1.28)$$

równanie (1.26) możemy zapisać w dobrze znanej postaci równania Fokkera–Plancka:

$$\frac{\partial P(m, t)}{\partial t} = \frac{\partial}{\partial m} [A(m)P(m, t)] + \frac{\partial^2}{\partial m^2} [B(m)P(m, t)]. \quad (1.29)$$

W dalszych rozważaniach poszukujemy równowagowego rozwiązania równania Fokkera–Plancka.

Rozwiązanie równowagowe równania Fokkera–Plancka

Równanie Fokkera–Plancka można przepisać w formie równania ciągłości dla gęstości prawdopodobieństwa:

$$\frac{\partial P(m, t)}{\partial t} = - \frac{\partial J(m, t)}{\partial m}, \quad (1.30)$$

gdzie $J(m, t) = -A(m)P(m, t) - \frac{\partial}{\partial m} [B(m)P(m, t)]$ oznacza strumień gęstości prawdopodobieństwa postaci:

$$J(m, t) = J_V(m, t) + J_D(m, t). \quad (1.31)$$

Składnik $J_V(m, t) = -A(m)P(m, t)$ wyraża konwekcyjny prąd prawdopodobieństwa, natomiast wyrażenie $J_D(m, t) = -\frac{\partial}{\partial m} [B(m)P(m, t)]$ jest dyfuzyjnym prądem gęstości prawdopodobieństwa.

⁶ W rzeczywistości wystarczyłoby całkowanie pomiędzy skończonymi granicami zmiany dochodu.

Ponieważ w naszych rozważaniach poszukujemy rozwiązań równowagowych (równowaga dynamiczna) przyjmujemy, że:

$$J(m, t) = 0, \quad \text{dla każdego } m. \quad (1.32)$$

Warunek ten oznacza, że interesują nas rozwiązania równowagowe odpowiadające sytuacji, w której wszystkie gospodarstwa domowe mają ustalone wielkości dochodów i nie następuje już ich zmiana w czasie. Zatem otrzymujemy:

$$A(m)P(m) = -\frac{\partial}{\partial m} [B(m)P(m)], \quad (1.33)$$

czyli

$$\frac{A(m)}{B(m)} = -\frac{\partial}{\partial m} \ln [B(m)P(m)]. \quad (1.34)$$

Ostatecznie, poszukiwane rozwiązanie równowagowe w reprezentacji Itô przybiera postać [6, 76]:

$$P(m) = \frac{\text{const}}{B(m)} \exp \left[-\int_{m_p}^m \frac{A(m')}{B(m')} dm' \right], \quad (1.35)$$

gdzie m_p jest minimalną wartością dochodu uzyskiwaną przez gospodarstwa domowe.

Otrzymane rozwiązanie wymaga komentarza. Po pierwsze, przeprowadzone rozumowanie jest pewnym uproszczeniem, które daje jednak poprawne wyniki. Gdybyśmy chcieli uzyskać równanie Fokkera–Plancka w postaci ścisłej a nie jako pewne przybliżenie, należałoby funkcje intensywności procesu stochastycznego $\Gamma(\cdot)$ uczynić zależnymi od pewnego małego parametru ϵ tak aby przy $\epsilon \rightarrow 0$ spełnione były założenia (1.24) oraz (1.25) [76]. Jednak w naszym podejściu nie dysponujemy tego typu parametrem.

Po drugie, w naszych rozważaniach przyjęliśmy, że skala czasowa jest rzędu jednego miesiąca, a nas interesuje zmienność roczna. Tak więc otrzymane rozwiązanie równowagowe rozumiemy jako rozwiązanie długookresowe w skali czasowej jednego roku.

Po trzecie, w powyższym wyprowadzeniu, pominęliśmy kwestię dylematu Itô–Stratonowicza. Ponieważ oba podejścia prowadzą do niemal identycznego rozwiązania równowagowego⁷, to w naszym wypadku wybieramy prostszą regułę in-

⁷ Chodzi tutaj o to, że w obu przypadkach czynniki wykładnicze są identyczne, a różnią się jedynie czynniki przedwykładnicze.

interpretacyjną Itô. Zatem, do współczynnika $C(m)$ podstawiamy wartość dochodu m przed skokiem. Równanie (1.1) interpretujemy wówczas następująco:

$$m(t + \Delta t) - m(t) = -A(m(t)) \Delta t + C(m(t)) \int_t^{t+\Delta t} \eta(t') dt'. \quad (1.36)$$

Równanie (1.35) posłuży nam do wyznaczenia gęstości prawdopodobieństw opisujących dochody każdej z klas społecznych, tj. gospodarstw domowych o niskich, średnich i wysokich dochodach a co najważniejsze do wyznaczenia ujednoliconej formuły gęstości prawdopodobieństwa opisującej dochody wszystkich trzech klas społecznych. Jest to zasadniczy cel teoretyczny niniejszej rozprawy doktorskiej. Jego realizacja (tutaj metodami fizyki statystycznej) stanowi oryginalny wkład do obecnego stanu wiedzy dotyczącej analizy dochodów gospodarstw domowych.

1.3 PRAWO BOLTZMANN-GIBBSA

W przypadku gospodarstw domowych o niskich dochodach możemy założyć, że zmiany ich dochodu Δm nie zależą od dochodu m uzyskiwanego dotychczas. Założenie to opiera się na obserwacji, że głównym źródłem ich dochodów są zarobki otrzymywane w postaci comiesięcznych pensji. Zatem, mamy do czynienia z addytywnym procesem stochastycznym, dla którego współczynniki $A(m)$ i $B(m)$ są stałe [2, 37, 74, 75]:

$$A(m) = A_0, \quad B(m) = B_0. \quad (1.37)$$

Wstawiając je do równania (1.35) otrzymujemy gęstość prawdopodobieństwa postaci⁸:

$$P(m) = \frac{\text{const}}{B_0} \exp \left(- \int_0^m \frac{A_0}{B_0} dm' \right) = \frac{\text{const}}{B_0} \exp \left(- \frac{A_0}{B_0} m \right) = C_0 \exp \left(- \frac{A_0}{B_0} m \right), \quad (1.38)$$

gdzie współczynnik $C_0 = \frac{\text{const}}{B_0}$ jest stałą normalizacyjną wyznaczoną z warunku:

$$\int_0^\infty P(m) dm = 1. \quad (1.39)$$

⁸ Jako dolną granicę całkowania we wzorze (1.35) wybieramy zero, ponieważ w praktyce dla danych empirycznych m_p jest bardzo małe, rzędu 1 euro lub 1 złotówki.

Wartość tego współczynnika zostanie wyznaczona w dalszej części.

Uzyskany wynik (1.38) jest identyczny z prawem Boltzmann–Gibbsa wywodzącym się z podstawowego modelu zderzeń [2, 6, 19].

Podstawowy model zderzeń

Podstawowy model zderzeń opiera się na analogii do zderzeń dwóch cząstek, podczas którego następuje zmiana energii kinetycznej każdej z nich. Istota tego modelu polega na tym, że w danej chwili dwa losowo wybrane gospodarstwa domowe dokonują pomiędzy sobą wymiany posiadanych zasobów finansowych⁹ zgodnie z określonymi regułami, nie biorąc pod uwagę wymiany dóbr i usług. Przyjmując, że proces wymiany jest niezależny od transakcji we wcześniejszych chwilach, dynamikę układu można opisać procesem Markowa:

$$\begin{pmatrix} m_i(t+1) \\ m_j(t+1) \end{pmatrix} = \begin{pmatrix} \epsilon_{ij}(t) & \epsilon_{ij}(t) \\ 1 - \epsilon_{ij}(t) & 1 - \epsilon_{ij}(t) \end{pmatrix} \begin{pmatrix} m_i(t) \\ m_j(t) \end{pmatrix}, \quad (1.40)$$

gdzie $m_i(t)$ oznacza dochód gospodarstwa domowego i w chwili t , a $0 \leq \epsilon_{ij}(t) \leq 1$ jest tutaj czynnikiem losowym, zmieniającym się wraz z upływem czasu, np. kolejną wymianą zasobów finansowych pomiędzy gospodarstwem i oraz j [6, 19, 22].

Dokładniej rzecz biorąc, podstawowy model zderzeń odnosi się do zamkniętego systemu ekonomicznego, w którym całkowita ilość pieniędzy M oraz liczba gospodarstw domowych N jest stała w czasie. Odpowiada to sytuacji, w której nie istnieje produkcja, a wszelka aktywność ekonomiczna ograniczona jest do wymiany zasobów finansowych opisywanej następującymi regułami:

– para gospodarstw domowych i oraz j wymienia posiadane przez siebie pieniądze w taki sposób, że łączna ich ilość pozostaje lokalnie zachowana, a więc:

$$\begin{aligned} m_i(t+1) &= m_i(t) + \Delta m, \\ m_j(t+1) &= m_j(t) - \Delta m, \end{aligned} \quad (1.41)$$

⁹ Zasoby finansowe posiadane przez gospodarstwo domowe oraz jego dochód są oczywiście różnymi pojęciami, jednak nie popełnimy błędu, jeśli przyjmiemy, że są one opisywane tym samym kształtem gęstości prawdopodobieństwa, gdyż ich dynamika stochastyczna rządzi się analogicznymi regułami [2, 37].

stąd jest spełniona lokalna zasada zachowania dochodów:

$$m_i(t) + m_j(t) = m_i(t+1) + m_j(t+1), \quad (1.42)$$

gdzie zgodnie z (1.40) ilość wymienionych pieniędzy $\Delta m = (\epsilon_{ij}(t) - 1)m_i(t) + \epsilon_{ij}(t)m_j(t)$;

– w wyniku zaistniałej wymiany żadne z gospodarstw domowych w chwili $t + 1$ nie może ponosić strat, co implikuje:

$$m_i(t) \geq 0. \quad (1.43)$$

Niech $P(m)$ oznacza gęstość prawdopodobieństwa dochodów zdefiniowaną w ten sposób, że $NP(m)dm$ jest liczbą gospodarstw domowych, które osiągają dochód pomiędzy m , a $m + dm$. Postać rozwiązania stacjonarnego w przypadku układu opisywanego równaniami (1.40–1.43) będzie odpowiadała stanowi równowagi termodynamicznej, w którym dochód m_i pojedynczego gospodarstwa domowego może nawet silnie fluktuować, jednak rozkład $P(m)$ odnoszący się do całego układu – społeczeństwa, nie ulega zmianie w czasie [6, 19].

Analityczną postać $P(m)$ można uzyskać stosując do naszego układu podejście fizyki statystycznej. Po pierwsze należy zauważyć, że zamiast o energii E_i cząstki i , mówimy o dochodzie m_i i -tego gospodarstwa domowego. Całkowita energia układu E będzie w tym wypadku reprezentowana przez całkowity dochód M wszystkich gospodarstw domowych. Ponieważ układ nasz jest izolowany (co wynika z założenia, że $M = \text{const}$) a liczba rozważanych gospodarstw domowych N jest stała w czasie, możemy zastosować do jego opisu zespół kanoniczny.

W tym celu należy podzielić rozważany układ na dwa podukłady, w których całkowite dochody wynoszą odpowiednio m_1 i m_2 a liczba gospodarstw domowych w podukładzie pierwszym $N_1 \ll N_2$. W dalszym ciągu interesuje nas tylko gęstość prawdopodobieństwa dochodów w podukładzie pierwszym. Drugi podukład traktujemy jak termostat („zbiornik cieplny”) zapewniający stałą „temperaturę”¹⁰ T w pierwszym podukładzie. Zatem, przeprowadzając rozumowanie analogiczne do tego, które stosuje się w przypadku układów termodynamicznych o ustalonej objętości, liczbie cząstek i temperaturze, pozostających w kontakcie z termostatem

¹⁰ Interpretacja wprowadzonej tutaj w sposób arbitralny temperatury zostanie podana w dalszej części rozważań.

możemy stwierdzić, że pierwszy podukład jest zespołem kanonicznym, w którym gęstość prawdopodobieństwa dochodów gospodarstw domowych przyjmuje postać¹¹ [83]:

$$P(m) = C_0 \exp\left(-\frac{m}{T}\right). \quad (1.44)$$

Stałą C_0 wyznaczamy z warunku normalizacyjnego (1.39):

$$\int_0^\infty P(m) dm = C_0 \int_0^\infty \exp\left(-\frac{m}{T}\right) dm = 1, \quad (1.45)$$

czyli:

$$C_0 = \frac{1}{T}. \quad (1.46)$$

Natomiast z warunku:

$$\langle m \rangle = \int_0^\infty m P(m) dm = \frac{M}{N}, \quad (1.47)$$

otrzymujemy, że:

$$\frac{1}{T} \int_0^\infty m \exp\left(-\frac{m}{T}\right) dm = \frac{M}{N}, \quad (1.48)$$

a stąd:

$$T = \frac{M}{N}. \quad (1.49)$$

Dzięki wyrażeniu (1.49) temperaturę możemy interpretować jako średni dochód pojedynczego gospodarstwa domowego. Zatem, gęstość rozkładu dochodów przyjmuje postać [6, 19]:

$$P(m) = \frac{1}{T} \exp\left(-\frac{m}{T}\right). \quad (1.50)$$

Porównując powyższy wynik z formułą (1.38), możemy traktować powyższy wynik jako ostateczną postać gęstości prawdopodobieństwa wyrażającego prawo Boltzmannna–Gibbsa, gdzie $T = \frac{B_0}{A_0}$ [2, 37, 73–75].

Dopełnienie dystrybucji $\Pi(m)$, czyli ułamek gospodarstw domowych uzyskujących dochód większy bądź równy m , wyznaczamy ze wzoru:

$$\Pi(m) = \int_m^\infty P(m') dm'. \quad (1.51)$$

Wykorzystując wyrażenie (1.50) otrzymujemy ostatecznie, że:

$$\Pi(m) = \int_m^\infty \frac{1}{T} \exp\left(-\frac{m'}{T}\right) dm' = \exp\left(-\frac{m}{T}\right). \quad (1.52)$$

¹¹ Dla uproszczenia przyjęto, że stała Boltzmannna $k = 1$.

W przypadku empirycznych dopełnień dystrybuant dotyczących rocznych dochodów gospodarstw domowych w Unii Europejskiej, rozkład ten dobrze opisuje niskie dochody gospodarstw domowych (patrz rozdział 4.4, rysunki 4.4 – 4.6).

Zakończenie dyskusji o podstawowym modelu zderzeń wymaga komentarza na temat założeń tego modelu, które stanowią pewną idealizację. Założenie, że mamy do czynienia z zamkniętym systemem ekonomicznym wynika z faktu, iż celem tego modelu jest (poprzez analogię do fizyki statystycznej) zaproponowanie najprostszego modelu ekonomicznego odtwarzającego obserwowane empiryczne dopełnienia dystrybuant. Z kolei, założenie o stałości całkowitej ilości pieniądza bierze się stąd, że gospodarstwa domowe same z siebie nie mogą zwiększać ilości pieniądza w gospodarce (mogą tylko produkować dobra). Jedynie bank centralny ma monopol na zmianę bazy monetarnej [84].

W przeciwieństwie do zwykłych podmiotów gospodarczych, bank centralny lub rząd może wprowadzić pieniądze do gospodarki, a tym samym dokonać zmiany całkowitej ilości pieniędzy w gospodarce. Proces ten jest analogiczny do napływu energii do układu ze źródeł zewnętrznych, np. Ziemia otrzymuje energię od Słońca. Opis takich zjawisk, fizycy zaczynają od rozważania układu zamkniętego, będącego w równowadze termicznej, a następnie uogólniają do układu otwartego, dopuszczając możliwość dopływu energii ze źródeł zewnętrznych. Dopóki tempo napływu pieniędzy z banku centralnego jest wolne w porównaniu z procesami relaksacji w gospodarce i nie powoduje hiperinflacji, układ znajduje się w kwazistacjonarnej równowadze statystycznej z wolno zmieniającymi się parametrami. Sytuacja ta jest analogiczna do powolnego ogrzewania czajnika, przy czym temperatura czajnika jest dobrze zdefiniowana w dowolnym momencie czasu, ale rośnie powoli. Pieniądz może również napływać do gospodarki poprzez międzynarodowe transfery finansowe. Proces ten opiera się na skomplikowanych zależnościach dotyczących współlistnienia na świecie wielu walut [85]. W modelach zderzeń stosuje się idealizację gospodarki zamkniętej ograniczonej do jednego kraju z jedną walutą. Taka idealizacja jest powszechnie przyjmowana w literaturze ekonomicznej (np. w książce [86] tylko ostatni – 23 rozdział zajmuje się problemami otwartej gospodarki).

Lokalne założenie o stałej ilości pieniądza w gospodarce wyrażone przez równanie (1.41) jest kluczowym elementem dla skutecznego funkcjonowania przepływu pieniądza w gospodarce. Jeśliby pozwolić gospodarstwom domowym na „produkcję” pieniędzy, to mogłyby one drukować pieniądze i kupować towary za nic, co prowadziłoby do katastrofy inflacyjnej.

Należy zauważyć, że makroekonomiczne kwestie polityki pieniężnej, takie jak podaż i popyt pieniądza [86] są pomijane w rozważaniach dotyczących podstawowego modelu zderzeń. Głównym celem tego modelu jest zaproponowanie mechanizmu odpowiedzialnego za pojawienie się określonego (w tym przypadku wykładniczego) rozkładu prawdopodobieństwa dochodów gospodarstw domowych. Aby to osiągnąć wprowadza się makroekonomiczne uproszczenia (opisane powyżej) prowadzące do zapewnienia stabilności całego układu i istnienia równowagi statystycznej w modelu. Pojęcie „równowagi” jest bardzo częstą idealizacją w literaturze ekonomicznej. W modelach zderzeń następuje rozszerzenie tej koncepcji do równowagi statystycznej, która charakteryzuje się stacjonarną gęstością rozkładu prawdopodobieństwa dochodów gospodarstw domowych, w przeciwieństwie do równowagi mechanicznej, w której „siły” popytu i podaży pieniądza równoważą się [2].

Przyjęcie powyższych założeń byłoby błędne w przypadku gospodarstw domowych o średnich i wysokich dochodach głównie dlatego, że inna jest dynamika uzyskiwania przez nie dochodu. Gospodarstwa te mogą, np. „kreować” pieniądza mając dostęp do międzynarodowych transferów finansowych (jako właściciele firm). Trudno też w tym przypadku mówić o stałej liczbie gospodarstw domowych, gdyż szczególnie gospodarstwa domowe o wysokich dochodach często zmieniają swoje miejsce zamieszkania. Zatem założenia podstawowego modelu zderzeń byłyby w tym przypadku zbyt grube.

Mając na uwadze fakt innej dynamiki zmian dochodu gospodarstw domowych o wysokich dochodach zaproponowałem w pracy model błędzenia losowego, mającego na celu wyjaśnienie mechanizmu odpowiedzialnego za pojawienie się słabego prawa Pareto w przypadku tych gospodarstw (patrz rozdział 2.1).

1.4 PRAWA PARETO

W przypadku gospodarstw domowych o średnich i dużych dochodach możemy założyć, że zmiana dochodu Δm jest proporcjonalna do dochodu uzyskiwanego przez nie dotychczas, co dotyczy zarówno tej części dochodu, która ma charakter systematyczny jak i tej o charakterze losowym (jest to tzw. prawo Efektów Proporcjonalnych Gibrata [5–10]). W takim przypadku źródłem systematycznych dochodów gospodarstw domowych są w głównej mierze zyski kapitałowe (np. lokaty, obligacje). Natomiast źródłem dochodów losowych są np. zwroty z różnego rodzaju inwestycji i okazjonalne premie. Zatem, mamy tutaj do czynienia z multiplikatywnym procesem stochastycznym scharakteryzowanym przez współczynniki $A(m)$ oraz $B(m)$ postaci [2, 37, 74, 75]:

$$A(m) = am, \quad B(m) = bm^2. \quad (1.53)$$

Korzystając z równania (1.35), otrzymujemy gęstość rozkładu prawdopodobieństwa¹²:

$$\begin{aligned} P(m) &= \frac{\text{const}}{bm^2} \exp\left(-\int_{m_0}^m \frac{am'}{bm'^2} dm'\right) = \frac{\text{const}}{bm^2} \exp\left(\frac{a}{b} \ln \frac{m_0}{m}\right) = \\ &= \frac{\text{const}}{bm^2} \left(\frac{m_0}{m}\right)^{a/b} = \frac{C_0}{m^{a/b+2}} = \frac{C_0}{m^{\alpha+1}}, \end{aligned} \quad (1.54)$$

gdzie współczynnik $C_0 = \frac{\text{const}}{b} m_0^{a/b}$ oraz wykładnik Pareto $\alpha = \frac{a}{b} + 1$. Korzystając z warunku normalizacyjnego postaci:

$$\int_{m_0}^{\infty} P(m) dm = 1, \quad (1.55)$$

otrzymujemy, że:

$$C_0 \int_{m_0}^{\infty} \frac{1}{m^{\alpha+1}} dm = 1, \quad (1.56)$$

skąd:

$$C_0 = \alpha m_0^{\alpha+1}. \quad (1.57)$$

¹² Tutaj za dolną granicę całkowania przyjmujemy parametr m_0 , którego interpretacja zostanie podana w dalszej części pracy.

Zatem, ostatecznie:

$$P(m) = \frac{\alpha}{m_0} \left(\frac{m_0}{m} \right)^{\alpha+1} \quad \text{dla } m \geq m_0, \quad (1.58)$$

co stanowi poszukiwaną zależność potęgową.

Silne prawo Pareto

Formułę (1.58) możemy zapisać w ogólniejszej postaci:

$$P(m) = \begin{cases} \frac{\alpha}{m_0} \left(\frac{m_0}{m} \right)^{\alpha+1} & \text{dla } m \geq m_0, \\ 0 & \text{dla } m < m_0. \end{cases} \quad (1.59)$$

Otrzymaliśmy w ten sposób gęstość prawdopodobieństwa reprezentującą silne prawo Pareto [3–6, 71–75]. Prawo to zostało sformułowane przez Vilfredo Pareto na podstawie analizy danych empirycznych dotyczących właśnie dochodów gospodarstw domowych.

Vilfredo Pareto jest pionierem badań w dziedzinie analizy dochodów ludności. Jednym z jego największych odkryć jest zaobserwowanie, że dopełnienia dystrybucji tych dochodów cechują się małą zmiennością w „przestrzeni” i czasie. Kształtem nie przypominają one rozkładów, które otrzymalibyśmy gdyby zdobywanie i posiadanie bogactwa miało charakter przypadkowy. Ponadto, rozkłady te są stabilne. Nawet gdybyśmy wykluczyli z procesu uzyskiwania dochodów najbogatszych, bądź najbiedniejszych członków społeczeństwa, to po pewnym okresie rozkład dochodów będzie ponownie przypominał kształtem początkową krzywą. Parametr m_0 występujący we wzorze (1.59) możemy interpretować jako graniczną, a zarazem najniższą wartość dochodu, jaką mogą osiągać gospodarstwa domowe. Gospodarstwa domowe uzyskujące niższe dochody nie byłyby w stanie przetrwać w społeczeństwie [3–6].

Istotną własnością rozkładu (1.59) jest rozbieżność momentów – pierwszego a stąd i wyższych rzędów dla $\alpha \leq 1$. Zatem, w takim przypadku zamiast mówić o wartościach oczekiwanych powinniśmy stosować kwantyle (które są zawsze skończone [87]). Na przykład, zamiast wartości średniej należałoby wprowadzić medianę (oczywiście przy analizie danych empirycznych estymatory momentów są zawsze skończone, ponieważ istnieje dochód o największej wartości obcinający rozkład Pareto).

Dodatkowo, w przypadku rozbieżności wartości średniej nie istnieje charakterystyczna skala dla dochodów uzyskiwanych przez najbogatsze gospodarstwa domowe (patrz rozdział 4.4, rysunki 4.10 – 4.11). Tak więc hierarchiczna, samopodobna (w sensie probabilistycznym) struktura klasy społecznej o najwyższych dochodach jest bezskalowa. Inaczej mówiąc, każde z gospodarstw domowych należące do tej klasy społecznej uzyskuje dochód opisywany analogiczną dynamiką stochastyczną [88].

Na podstawie definicji (1.51) oraz wyrażenia (1.59) dopełnienie dystrybuanty przyjmuje postać:

$$\Pi(m) = \begin{cases} \left(\frac{m_0}{m}\right)^\alpha & \text{dla } m \geq m_0, \\ 1 & \text{dla } m < m_0. \end{cases} \quad (1.60)$$

Wartość 1 dla $m < m_0$ jest oczywista i wynika z faktu, że zgodnie z przyjętą interpretacją m_0 wszystkie gospodarstwa domowe osiągają dochód większy niż m_0 . Wartość wykładnika Pareto α wyznacza się na podstawie danych empirycznych. W przypadku, gdy $\alpha = 3/2$ mówimy o najsilniejszym prawie Pareto. Jest to średnia wartość wykładnika α jaką otrzymał Pareto w swoich badaniach [4–6].

Słabe prawo Pareto

Współczesne badania wykazały, że silne prawo Pareto, stoi częściowo w sprzeczności z obecnymi danymi empirycznymi, a średnia wartość wykładnika α jest bliższa 2 [5, 6]. Pozostaje ono jednak zawsze spełnione dla dużych wartości m (tzn. $m \gg m_0$), co można wyrazić poprzez słabe prawo Pareto:

$$\Pi(m) \sim (m/m_0)^{-\alpha}. \quad (1.61)$$

Jak widać słabe prawo Pareto ma charakter asymptotyczny. Ostatecznie:

$$\frac{\Pi(m)}{(m/m_0)^{-\alpha}} \rightarrow 1, \quad \text{jeśli } m \rightarrow \infty \quad (1.62)$$

lub równoważnie:

$$\Pi(m) = [1 + \epsilon(m)](m/m_0)^{-\alpha}, \quad \text{gdzie } \epsilon(m) \rightarrow 0, \quad \text{jeśli } m \rightarrow \infty. \quad (1.63)$$

Wolnomalejąca funkcja $\epsilon(m)$ jest poprawką do prawa skalowania (1.61). Podobnie jak poprzednio, wartość α wyznacza się na podstawie analizy danych empirycznych [4–6].

Badaniem własności słabego rozkładu Pareto (wzór (1.61)) zajmował się m.in. Benoit Mandelbrot (patrz dodatek A).

W przypadku empirycznych dopełnień dystrybuant dotyczących rocznych dochodów gospodarstw domowych w Polsce oraz Unii Europejskiej, słabe prawo Pareto dobrze opisuje gospodarstwa domowe o średnich oraz wysokich dochodach (patrz rozdział 3.3 oraz 4.4, rysunki 3.7 – 3.9, 3.13 – 3.18 oraz 4.7 – 4.11).

1.5 MODEL YAKOVENKO

Dotychczas przedstawiałem modele pomocnicze, opisujące z osobna klasy społeczne. Modele te posłużą do wstępnego wyznaczenia, w oparciu o dane empiryczne, niezbędnych parametrów. **Jednakże, zasadniczym punktem wyjścia niniejszej rozprawy jest model Yakovenko, opisujący dwie klasy społeczne, tj. gospodarstwa domowe o niskich i średnich dochodach. Właśnie ten model zostanie w pracy rozszerzony na trzy klasy społeczne. Rozszerzenie to stanowi kluczowy element teoretyczny niniejszej rozprawy doktorskiej, umożliwiając narreszcie traktowanie społeczeństwa jako całości.**

W modelu Yakovenko [2, 37, 74, 75] dopuszczamy współistnienie procesu addytywnego i moltiplikatywnego a zatem zakładamy, że gospodarstwo domowe może uzyskiwać dochody zarówno w postaci comiesięcznych pensji jak i np. zysków kapitałowych. Z sytuacją taką mamy do czynienia również wtedy, gdy gospodarstwo domowe uzyskuje procentowy wzrost dochodu (np. w wyniku lokaty bankowej – jest to wówczas proces moltiplikatywny) oraz, gdy gospodarstwo domowe otrzymuje podwyżkę (proces addytywny). Mając na uwadze powyższe możemy przyjąć, że:

$$A(m) = A_0 + am, \quad B(m) = B_0 + bm^2 = b(m_0^2 + m^2), \quad (1.64)$$

gdzie wprowadziliśmy oznaczenie¹³ $B_0 = bm_0^2$ oraz stałe A_0 i B_0 są dodatnie.

¹³ Pojawiającego się tutaj parametru m_0 nie należy mylić z parametrem m_0 występującym w słabym prawie Pareto.

Korzystając z równania (1.35) możemy wyznaczyć gęstość prawdopodobieństwa¹⁴:

$$\begin{aligned}
 P(m) &= \frac{\text{const}}{b(m_0^2 + m^2)} \exp \left(- \int_0^m \frac{A_0 + am'}{b(m_0^2 + m'^2)} dm' \right) = \\
 &= \frac{\text{const}}{b(m_0^2 + m^2)} \exp \left[- \left(\frac{A_0}{b} \int_0^m \frac{1}{m_0^2 + m'^2} dm' + \frac{a}{b} \int_0^m \frac{m'}{m_0^2 + m'^2} dm' \right) \right] = \\
 &= \frac{\text{const}}{b(m_0^2 + m^2)} \exp \left[- \frac{m_0 A_0}{b m_0^2} \arctan \frac{m}{m_0} - \ln \left(\frac{m_0^2 + m^2}{m_0^2} \right)^{a/2b} \right] = \\
 &= \frac{\text{const}}{b(m_0^2 + m^2)} \exp \left[- \frac{m_0 A_0}{B_0} \arctan \frac{m}{m_0} - \ln \left(\frac{m_0^2 + m^2}{m_0^2} \right)^{a/2b} \right] = \\
 &= \frac{\text{const}}{b m_0^2} \frac{\exp \left(- \frac{m_0}{T} \arctan \frac{m}{m_0} \right)}{\left[1 + \left(\frac{m}{m_0} \right)^2 \right]^{a/2b}} = \\
 &= C' \frac{\exp \left(- \frac{m_0}{T} \arctan \frac{m}{m_0} \right)}{\left[1 + \left(\frac{m}{m_0} \right)^2 \right]^{\frac{\alpha+1}{2}}}, \tag{1.65}
 \end{aligned}$$

gdzie $T = \frac{B_0}{A_0}$, $\frac{\alpha+1}{2} = \frac{a}{2b} + 1$ oraz $C' = \frac{\text{const}}{b m_0^2}$.

Rozkład (1.65) jest połączeniem prawa wykładniczego z prawem potęgowym. Dla $m \ll m_0$ przechodzi on w wykładnicze prawo Boltzmanna–Gibbsa (dominuje wówczas addytywny proces stochastyczny), co wynika z rozwinięcia $\arctan(\dots)$ w szereg Taylora:

$$\arctan \frac{m}{m_0} = \frac{m}{m_0} - \frac{1}{3} \left(\frac{m}{m_0} \right)^3 + \vartheta \left(\left(\frac{m}{m_0} \right)^5 \right). \tag{1.66}$$

Wtedy, $P(m) \sim \exp \left(- \frac{m}{T} \right)$.

Dla $m \gg m_0$ rozkład (1.65) staje się słabym prawem Pareto (dominuje multiplikatywny proces stochastyczny), gdyż wtedy $\arctan \frac{m}{m_0} \approx \frac{\pi}{2}$. Tak więc, $P(m) \sim \left(\frac{m}{m_0} \right)^{-(\alpha+1)}$.

Dla $m \approx m_0$ wkłady do $B(m)$ pochodzące od procesu addytywnego i multiplikatywnego są istotne. Parametr m_0 można traktować jako umowną granicę pomiędzy prawem Boltzmanna–Gibbsa a słabym prawem Pareto, czyli pomiędzy procesem addytywnym a multiplikatywnym. Z kolei, parametr T można uważać za

¹⁴ Podobnie jak poprzednio, za dolną granicę całkowania we wzorze (1.35) wybieramy zero, ponieważ w praktyce dla danych empirycznych m_p jest bardzo małe, rzędu 1 euro lub 1 złotówki.

(zdefiniowaną w rozdziale 1.3) temperaturę, natomiast α jest wykładnikiem Pareto wprowadzonym w rozdziale 1.4 [2, 37, 74, 75].

Zauważmy, że dla rozkładu danego wzorem (1.65) nie daje się niestety wyznaczyć dopełnienia dystrybuanty $\Pi(m)$ w postaci analitycznej.

1.6 PRAWO EFEKTÓW PROPORCJONALNYCH

W kolejnych dwóch podrozdziałach omawiane są modele, które dobrze opisują niskie i średnie dochody gospodarstw domowych w Polsce. Jednym z nich jest Prawo Efektów Proporcjonalnych Gibrata [5–10]. W modelu tym zakłada się, że współczynnik dryfu $A(m)$ i współczynnik $B(m)$ związany z dyfuzją, przyjmują postać [2, 71, 72]:

$$A(m) = am, \quad B(m) = bm^2. \quad (1.67)$$

Dokonując prostej zamiany zmiennych w równaniu Fokkera–Plancka (1.29) (tzn. m na $\ln m$) otrzymujemy:

$$\begin{aligned} \frac{\partial P(m, t)}{\partial t} &= \frac{d \ln m}{dm} \frac{\partial}{\partial \ln m} [A(m)P(m, t)] + \frac{\partial}{\partial m} \left\{ \frac{d \ln m}{dm} \frac{\partial}{\partial \ln m} [B(m)P(m, t)] \right\} = \\ &= \frac{1}{m} \frac{\partial}{\partial \ln m} [A(m)P(m, t)] + \frac{\partial}{\partial m} \left\{ \frac{1}{m} \frac{\partial}{\partial \ln m} [B(m)P(m, t)] \right\} = \\ &= \frac{1}{m} \frac{\partial}{\partial \ln m} [A(m)P(m, t)] - \frac{1}{m^2} \frac{\partial}{\partial \ln m} [B(m)P(m, t)] + \\ &\quad + \frac{1}{m} \frac{d \ln m}{dm} \frac{\partial}{\partial \ln m} \left\{ \frac{\partial}{\partial \ln m} [B(m)P(m, t)] \right\} = \\ &= \frac{1}{m} \frac{\partial}{\partial \ln m} [A(m)P(m, t)] - \frac{1}{m^2} \frac{\partial}{\partial \ln m} [B(m)P(m, t)] + \\ &\quad + \frac{1}{m^2} \frac{\partial^2}{\partial (\ln m)^2} [B(m)P(m, t)]. \end{aligned} \quad (1.68)$$

Mnożąc ostatnie wyrażenie stronami przez m i wstawiając współczynniki $A(m)$ i $B(m)$ w postaci (1.67) możemy napisać, że:

$$\begin{aligned} \frac{\partial}{\partial t} [mP(m, t)] &= \frac{\partial}{\partial \ln m} [amP(m, t)] - \frac{1}{m} \frac{\partial}{\partial \ln m} [bm^2P(m, t)] + \\ &\quad + \frac{1}{m} \frac{\partial^2}{\partial (\ln m)^2} [bm^2P(m, t)]. \end{aligned} \quad (1.69)$$

Następnie podstawiamy:

$$x = \ln m, \quad dx = \frac{1}{m} dm, \\ Q(x, t) = P(m, t) \left| \frac{dx}{dm} \right|^{-1} = mP(m, t), \quad (1.70)$$

uzyskując:

$$\frac{\partial Q(x, t)}{\partial t} = a \frac{\partial Q(x, t)}{\partial x} - \frac{b}{e^x} \frac{\partial}{\partial x} [e^x Q(x, t)] + \frac{b}{e^x} \frac{\partial^2}{\partial x^2} [e^x Q(x, t)]. \quad (1.71)$$

Stąd, obliczając poszczególne pochodne i dokonując redukcji wyrazów podobnych, otrzymujemy:

$$\frac{\partial Q(x, t)}{\partial t} = -(-a - b) \frac{\partial Q(x, t)}{\partial x} + b \frac{\partial^2 Q(x, t)}{\partial x^2}. \quad (1.72)$$

Rozwiązaniem powyższego równania jest:

$$Q(x, t) = \frac{1}{\sqrt{4\pi bt}} \exp \left\{ -\frac{[x - (-a - b)t]^2}{4bt} \right\}. \quad (1.73)$$

Jest to rozkład Gaussa o średniej $\mu_t = (-a - b)t$ i wariancji $\sigma_t^2 = 2bt$. Jego szerokość (czyli $2\sigma_t$) wzrasta proporcjonalnie do pierwiastka kwadratowego z czasu zatem, wyrażenie (1.73) stanowi rozwiązanie niestacjonarne równania Fokkera–Plancka, opisujące „rozpływanie się” rozkładu Gaussa w zmiennej x .

Współczynniki $A(m)$ i $B(m)$ zdefiniowane w (1.67) są identyczne jak w przypadku słabego prawa Pareto. Zasadnicza różnica między prawem Efektów Proporcjonalnych a słabym prawem Pareto polega na tym, iż w pierwszym przypadku zakładamy, że zmiany dochodów gospodarstw domowych ze wszystkich grup społecznych są proporcjonalne do dochodów uzyskiwanych przez nie dotychczas. Natomiast w drugim przypadku słabe prawo Pareto dopuszcza takie założenie, ale tylko dla gospodarstw domowych uzyskujących dostatecznie duże dochody (czyli dochody średnie i wysokie). Dzięki narzuceniu ograniczenia wielkości dochodu z dołu możemy uzyskać, w ramach multiplikatywnego procesu stochastycznego, rozwiązanie równowagowe będące właśnie słabym prawem Pareto [2] (patrz rozdział 1.4).

Jeśli w rozwiązaniu (1.73) powrócimy do zmiennej m , to korzystając z podstawień (1.70), otrzymujemy rozkład logarytmiczno–normalny (log–normalny), który

dobrze opisuje niskie dochody uzyskiwane przez gospodarstwa domowe w Polsce (patrz rozdział 3.3). Zatem [2,6,82]:

$$P(m, t) = \frac{1}{\sqrt{2\pi\sigma_t^2 m}} \exp \left[-\frac{(\ln m - \mu_t)^2}{2\sigma_t^2} \right] \quad (1.74)$$

natomiast, dopełnienie dystrybuanty wyraża się wzorem [6,71,72]:

$$\begin{aligned} \Pi(m) &= \int_m^\infty \frac{1}{\sqrt{2\pi\sigma_t^2 m'}} \exp \left[-\frac{(\ln m' - \mu_t)^2}{2\sigma_t^2} \right] dm' = \\ &= \frac{1}{\sqrt{\pi}} \int_t^\infty e^{-t'^2} dt' = \frac{1}{2} \left(\frac{2}{\sqrt{\pi}} \int_t^\infty e^{-t'^2} dt' \right) = \\ &= \frac{1}{2} \left(1 - \frac{2}{\sqrt{\pi}} \int_0^t e^{-t'^2} dt' \right) = \frac{1}{2} (1 - \operatorname{erf}(t)), \end{aligned} \quad (1.75)$$

gdzie po drodze dokonano następującej zamiany zmiennych: $t' = \frac{\ln m' - \mu_t}{\sqrt{2}\sigma_t}$, $dt' = \frac{1}{\sqrt{2}\sigma_t m'} dm'$; jak zwykle $\operatorname{erf}(t) = \frac{2}{\sqrt{\pi}} \int_0^t e^{-t'^2} dt'$ jest funkcją błędu. Ostatecznie otrzymujemy:

$$\Pi(m) = \frac{1}{2} \left[1 - \operatorname{erf} \left(\frac{\ln m - \mu_t}{\sqrt{2}\sigma_t} \right) \right]. \quad (1.76)$$

W przypadku empirycznych dopełnień dystrybuant dotyczących rocznych dochodów uzyskiwanych przez gospodarstwa domowe w Polsce, rozkład ten dobrze opisuje gospodarstwa domowe o niskich dochodach (patrz rozdział 3.3, rysunki 3.4 – 3.6).

Należy zaznaczyć, że problemem otwartym pozostaje uzyskanie jednolitej formuły gęstości prawdopodobieństwa opisującej wszystkie trzy klasy społeczne, gdzie klasa społeczna o niskich dochodach byłaby opisywana rozkładem log-normalnym.

1.7 UOGÓLNIONY MODEL LOTKA–VOLTERRA

Kolejnym modelem, który dobrze opisuje niskie i średnie dochody gospodarstw domowych w Polsce jest uogólniony model Lotka–Volterra zaproponowany przez Solomona i Richmonda [2,5,6,38–41,71–73]. Zakłada on, że dochód gospodarstwa domowego zmienia się w czasie zgodnie z następującym równaniem różnicowym:

$$m_i(t + \Delta t) - m_i(t) = [r_i(t) - 1] \Delta t m_i(t) + a \Delta t m(t) - c(m_1, \dots, m_N, t) \Delta t m_i(t), \quad (1.77)$$

gdzie:

- $m_i(t)$ oznacza dochód i -tego gospodarstwa domowego w chwili t ;
- $r_i(t)$ jest białym szumem, którego własności

$$\text{brak dryfu: } \langle r_i(t) \rangle = 0, \quad (1.78)$$

$$\text{oraz skończoną wariancję: } \langle r_i(t)^2 \rangle = D \quad (1.79)$$

wykorzystujemy w dalszej części, przy czym na mocy założenia, wariancja D jest rzędu Δt . Wielkości $r_i(t)$ wyrażają losową stopę zwrotu z kapitału $m_i(t)$ w przedziale czasu $t, t + \Delta t$;

- $m(t)$ oznacza średni dochód w społeczeństwie (składającym się z N gospodarstw domowych) w chwili t :

$$m(t) = \frac{1}{N} \sum_j m_j(t); \quad (1.80)$$

- składnik $a\Delta t m(t)$ reprezentuje dochody gospodarstw domowych pochodzące z subsydów, świadczonych usług lub pomocy społecznej, stąd też jest on proporcjonalny do średniego dochodu $m(t)$. Składnik $a\Delta t m(t)$ gwarantuje, iż dochód gospodarstwa domowego nie spadnie poniżej pewnej wartości minimalnej a zatem, zawsze będzie większy od zera;

- współczynnik $c(m_1, \dots, m_N, t)$ eliminuje możliwość nieograniczonego wzrostu dochodów poszczególnych gospodarstw domowych. Wyraża on zewnętrzne ograniczenia takie jak, skończona ilość surowców i pieniędzy w gospodarce, wojny, katastrofy, a także uwzględnia procesy wolnego rynku: konkurencję pomiędzy inwestorami, postęp technologiczny, zmienne ceny różnego rodzaju aktywów, itp. Tym samym gwarantuje on, że wzrost średniego dochodu w społeczeństwie będzie następował do poziomu, który jest możliwy do osiągnięcia w obliczu panujących warunków zewnętrznych i dostępnych zasobów. W ogólności, współczynnik $c(m_1, \dots, m_N, t)$ opisuje stan gospodarki. Okresy, w których $-c(m_1, \dots, m_N, t)$ jest duże i dodatnie odpowiadają boomowi gospodarczemu, gdy średnie dochody w społeczeństwie wykazują tendencję wzrostową. Z kolei ujemne wartości $-c(m_1, \dots, m_N, t)$ są związane z recesją, gdy przeważnie mamy wtedy do czynienia z małymi bądź ujemnymi stopami zwrotu.

Można wykazać [6, 38–40], że równanie (1.77) nie posiada rozwiązania równowagowego, jednakże daje się je sprowadzić do równania na dochód względny $x_i(t) = \frac{m_i(t)}{m(t)}$, które posiada już rozwiązanie równowagowe. W tym celu wyrażenie (1.77) należy przetransformować do postaci niezależnej od współczynnika c (patrz dodatek C). Otrzymujemy, że [6, 38–40]:

$$\frac{dx_i(t)}{dt} = a - ax_i(t) + r_i(t)x_i(t). \quad (1.81)$$

Jest to nieliniowe równanie Langevina dynamiki stochastycznej z siłą oporu $a - ax_i(t)$ oraz siłą stochastyczną $r_i(t)x_i(t)$. Oczywiście, jest ono równoważne równaniu (1.2), gdy:

$$A(x_i) = ax_i - a, \quad C(x_i) = x_i. \quad (1.82)$$

W dalszym ciągu zmienne x_i można traktować jako zmienne statystycznie niezależne w przeciwieństwie do zmiennych m_i . Jak pokazaliśmy w rozdziale 1.2, z równania Langevina wynika równanie Fokkera–Plancka na gęstość prawdopodobieństwa $P(x_i, t)$. Równowagowe rozwiązanie tego równania przyjmuje, w naszych oznaczeniach, postać:

$$P(x_i) = \frac{\text{const}}{B(x_i)} \exp \left[- \int_{x_i(0)}^{x_i} \frac{A(x'_i)}{B(x'_i)} dx'_i \right], \quad (1.83)$$

gdzie $B(x'_i) = \frac{D}{2} C^2(x'_i)$ oraz $x_i(0) = \frac{m_{\min}}{m(0)}$, przy czym m_{\min} jest minimalną wartością dochodu (większą od zera).

Korzystając z równań (1.82), otrzymujemy rozwiązanie (1.83) w jawnej postaci:

$$P(x) = \frac{C'}{x^2} \exp \left[\frac{2}{D} \int_{x(0)}^x \frac{a(1-x')}{x'^2} dx' \right], \quad (1.84)$$

gdzie $C' = \frac{2\text{const}}{D}$ a indeks i został już (dla uproszczenia) opuszczony. Po obliczeniu powyższej całki otrzymujemy:

$$P(x) = C_0 \exp \left(- \frac{2a}{Dx} \right) x^{-\frac{2a}{D}-2}, \quad (1.85)$$

gdzie $C_0 = C' \exp \left(\frac{2a}{Dx_0} \right) x_0^{\frac{2a}{D}}$. Stosując podstawienie $\alpha = 1 + \frac{2a}{D}$ dostajemy gęstość prawdopodobieństwa dla względnego dochodu jednostek w uogólnionym modelu Lotka–Volterra:

$$P(x) = C_0 \frac{\exp \left(- \frac{\alpha-1}{x} \right)}{x^{1+\alpha}}. \quad (1.86)$$

Stałą C_0 można wyznaczyć z warunku normalizacyjnego:

$$\int_0^{\infty} P(x) dx = 1. \quad (1.87)$$

Zatem:

$$\int_0^{\infty} C_0 \frac{\exp\left(-\frac{\alpha-1}{x}\right)}{x^{1+\alpha}} dx = \frac{C_0}{(\alpha-1)^\alpha} \int_0^{\infty} e^{-t} t^{\alpha-1} dt = C_0 \frac{\Gamma(\alpha)}{(\alpha-1)^\alpha} = 1, \quad (1.88)$$

a stąd:

$$C_0 = \frac{(\alpha-1)^\alpha}{\Gamma(\alpha)}, \quad (1.89)$$

gdzie skorzystano z następującego podstawienia: $t = \frac{\alpha-1}{x}$; $dx = -\frac{x^2}{\alpha-1} dt$.

Ostatecznie gęstość rozkładu prawdopodobieństwa dla względnego dochodu gospodarstw domowych w uogólnionym modelu Lotka–Volterra przyjmuje postać odwrotnego rozkładu gamma [2, 5, 6, 38–41]:

$$P(x) = \frac{(\alpha-1)^\alpha}{\Gamma(\alpha)} \frac{\exp\left(-\frac{\alpha-1}{x}\right)}{x^{1+\alpha}}. \quad (1.90)$$

Powyższy wynik implikuje fakt, że względne dochody w społeczeństwie opisywane są przez rozkład równowagowy charakteryzujący się niezależnym od czasu wykładnikiem α . Jego wartość jest uwarunkowana jedynie stosunkiem parametru opisującego redystrybucję dochodów w społeczeństwie (współczynnik α) do zmienności D losowej stopy zwrotu z posiadanego przez gospodarstwo domowe kapitału. Wykładnik ten zbliża się do 1 w przypadku najbogatszych gospodarstw domowych, gdyż wówczas rośnie ryzyko (nawet znacząco) mierzone wariancją stopy zwrotu D [89].

Możemy zauważyć, że rozważana gęstość prawdopodobieństwa gwałtownie maleje do zera, gdy $x \rightarrow 0$, natomiast wolno zanika, gdy $x \rightarrow \infty$. Co więcej, to gwałtowne zanikanie do zera ma miejsce dla x tylko nieznacznie mniejszego od x^{\max} . Oznacza to, że wartość x^{\max} , dla której funkcja $P(x)$ osiąga maksimum, będzie z dobrym przybliżeniem efektywną wartością minimalną względnego dochodu, tzn. $x_{\min} \approx x^{\max}$.

Z warunku na ekstremum $\frac{dP(x)}{dx} = 0$, otrzymujemy:

$$x_{\min} = \frac{\alpha-1}{\alpha+1} \quad (1.91)$$

lub

$$\alpha = \frac{1 + x_{\min}}{1 - x_{\min}} = \frac{1 + m_{\min}/m}{1 - m_{\min}/m}. \quad (1.92)$$

W świetle wyrażenia (1.92) jest oczywiste, że wartość wykładnika α jest determinowana także przez stosunek minimalnego dochodu do średniego dochodu w społeczeństwie, który to stosunek jest niezależny od czasu t [6, 38–40].

Dopełnienie dystrybuanty gęstości (1.90) wyraża się wzorem:

$$\begin{aligned} \Pi(x) &= \int_x^\infty \frac{(\alpha - 1)^\alpha}{\Gamma(\alpha)} \frac{\exp\left(-\frac{\alpha-1}{x'}\right)}{x'^{1+\alpha}} dx' = \frac{1}{\Gamma(\alpha)} \int_0^t e^{-t'} t'^{\alpha-1} dt' = \\ &= 1 - \frac{1}{\Gamma(\alpha)} \int_t^\infty e^{-t'} t'^{\alpha-1} dt' = 1 - \frac{\Gamma(\alpha, t)}{\Gamma(\alpha)}, \end{aligned} \quad (1.93)$$

gdzie zastosowano następującą zamianę zmiennych: $t' = \frac{\alpha-1}{x'}$; $dx' = -\frac{x'^2}{\alpha-1} dt'$. Funkcja $\Gamma(\alpha, t) = \int_t^\infty e^{-t'} t'^{\alpha-1} dt'$ jest górną niekompletną funkcją gamma. Ostatecznie [5, 6, 71–73]:

$$\Pi(x) = 1 - \frac{\Gamma(\alpha, \frac{\alpha-1}{x})}{\Gamma(\alpha)}. \quad (1.94)$$

Rozkład ten dobrze opisuje względne, niskie i średnie dochody uzyskiwane przez polskie gospodarstwa domowe (patrz rozdział 3.3, rysunki 3.10 – 3.12).

MODELE DYNAMIKI DOCHODÓW ROZWINIĘTE W NINIEJSZEJ ROZPRAWIE

*„Cała nasza nauka, w porównaniu z rzeczywistością,
jest prymitywna i dziecinna
- ale nadal jest to najcenniejsza rzecz, jaką posiadamy.”*

ALBERT EINSTEIN

2.1 PRAWO PARETO – MODEL DRABINOWY

W niniejszym podrozdziale przedstawiam autorski model, będący uogólnieniem i uzupełnieniem rozważań zawartych w pracy [90], mający na celu opisanie przykładowego mechanizmu odpowiedzialnego za pojawienie się słabego prawa Pareto w przypadku gospodarstw domowych o wysokich dochodach.

Podzielmy gospodarstwa domowe o wysokich dochodach¹ na podgrupy. Do danej podgrupy o numerze $i = 1, 2, 3, \dots$, zaliczamy gospodarstwa domowe o względnych dochodach $\frac{m}{m_{\min}}$ zawartych w przedziale:

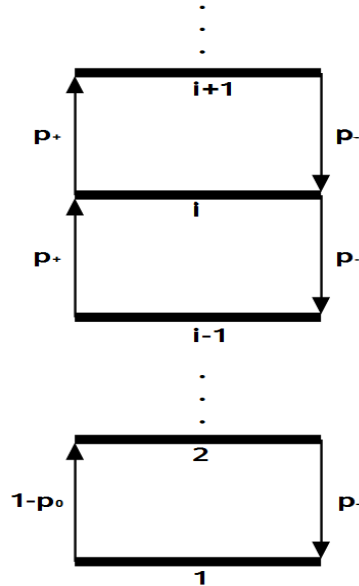
$$\begin{aligned} \frac{m_i}{m_{\min}} = q^{i-1} \leq \frac{m}{m_{\min}} < q^i = \frac{m_{i+1}}{m_{\min}} &\Leftrightarrow \\ \Leftrightarrow \ln \frac{m_i}{m_{\min}} = (i-1) \ln q \leq \ln \frac{m}{m_{\min}} < i \ln q = \ln \frac{m_{i+1}}{m_{\min}}, &\quad (2.1) \end{aligned}$$

gdzie m_{\min} jest minimalną wartością dochodu uzyskiwanego wśród wszystkich gospodarstw domowych o wysokich dochodach, natomiast parametr q definiuje podstawę strukturalnego podziału względnych dochodów gospodarstw domowych

¹ Jak się dalej okaże, tylko dla tego typu gospodarstw domowych daje się zastosować model drabinowy.

o wysokich dochodach. Dodatkowo zakładamy, że z roku na rok liczba gospodarstw domowych należących do poszczególnych podgrup zmienia się zgodnie z następującymi regułami (patrz rysunek 2.1):

- gospodarstwo domowe z danej podgrupy i , w następnym roku może znaleźć się jedynie w podgrupie $i + 1$, bądź podgrupie $i - 1$;
- prawdopodobieństwa przejść na jednostkę czasu wynoszą odpowiednio p_+ w przypadku przejść $i \rightarrow i + 1$ oraz p_- w przypadku przejść $i \rightarrow i - 1$ oraz $p_+ + p_- = 1$;
- dla najniższej podgrupy $i = 1$ wprowadzamy prawdopodobieństwo przetrwania, p_0 , w tej grupie, zatem $1 - p_0$ pełni tutaj rolę prawdopodobieństwa przejścia p_+ .



Rysunek 2.1: Schematycznie przedstawione błędzenie losowe gospodarstw domowych na „drabini” dochodów.

Niech $N_i(t)$ oznacza liczbę gospodarstw domowych należących do podgrupy i w chwili t (t oznacza tutaj np. miesiąc). Wówczas równanie dynamiki liczby gospodarstw domowych N_i opisane jest procesem Markowa:

$$N_i(t) = N_{i+1}(t-1)p_- + N_{i-1}(t-1)p_+, \quad i \geq 2, \quad (2.2)$$

$$N_1(t) = N_2(t-1)p_- + N_1(t-1)p_0. \quad (2.3)$$

Na razie skupimy się na, wynikającym z powyższego, podejściu statycznym.

Podejście statyczne

Ponieważ interesuje nas rozwiązanie równowagowe, to w zgodzie z równaniami (2.2) i (2.3) możemy wprowadzić warunki równowagi szczegółowej postaci:

$$N_i p_+ = N_{i+1} p_-, \quad i \geq 2, \quad (2.4)$$

$$N_2 p_- = (1 - p_0) N_1, \quad (2.5)$$

gdzie oczywiście N_i , $i = 1, 2, \dots$, jest już wielkością niezależną od czasu. Przekształcając równania (2.4) i (2.5) otrzymujemy:

$$N_{i+1} = q^{-\alpha} N_i, \quad i \geq 2, \quad (2.6)$$

$$N_2 = q_0^{-\alpha} N_1, \quad (2.7)$$

gdzie:

$$q^{-\alpha} = \frac{p_+}{p_-} \Leftrightarrow \alpha = \frac{\ln \frac{p_-}{p_+}}{\ln q} \quad (2.8)$$

a stąd wprowadziliśmy:

$$q_0^{-\alpha} = \frac{1 - p_0}{p_-} \Leftrightarrow \ln q_0 = \frac{\ln \frac{p_-}{1 - p_0}}{\alpha}. \quad (2.9)$$

Rozwiązując rekurencyjnie równanie (2.6) ostatecznie uzyskujemy:

$$N_i = (q^{-\alpha})^{i-2} q_0^{-\alpha} N_1, \quad i \geq 2, \quad (2.10)$$

$$N_2 = q_0^{-\alpha} N_1. \quad (2.11)$$

Korzystając z równań (2.10) i (2.11) możemy wyznaczyć równowagowe prawdopodobieństwo P_i tego, że losowo wybrane gospodarstwo domowe o wysokich dochodach należy do podgrupy i :

$$P_i = \frac{N_i}{N} = \frac{q_0^{-\alpha}}{1 + \frac{q_0^{-\alpha}}{1 - q^{-\alpha}}} \frac{1}{(q^\alpha)^{i-2}}, \quad i \geq 2, \quad (2.12)$$

$$P_1 = \frac{N_1}{N} = \frac{1}{1 + \frac{q_0^{-\alpha}}{1 - q^{-\alpha}}}, \quad (2.13)$$

gdzie N jest liczbą gospodarstw domowych o wysokich dochodach we wszystkich podgrupach:

$$N = N_1 + \sum_{i=2}^{\infty} N_i = N_1 + N_1 q_0^{-\alpha} \sum_{i=2}^{\infty} (q^{-\alpha})^{i-2} = N_1 \left(1 + \frac{q_0^{-\alpha}}{1 - q^{-\alpha}} \right). \quad (2.14)$$

Oczywiście spełniony jest warunek normalizacji:

$$P_1 + \sum_{i=2}^{\infty} P_i = 1. \quad (2.15)$$

Przynależność określonego gospodarstwa domowego do danej podgrupy i jest determinowana uzyskiwanym przez nie dochodem m . Zatem, wykorzystując klasyfikację (2.1), możemy zapisać prawdopodobieństwo P_i w postaci:

$$P_i = P(m_{\min} q^{i-1} \leq m < m_{\min} q^i) = P(m_i \leq m < m_{i+1}). \quad (2.16)$$

Z kolei prawdopodobieństwo, że losowo wybrane gospodarstwo domowe należy do podgrupy $j \geq i$ (inaczej mówiąc, że jego dochody są równe bądź przekraczają wartość $m_i = m_{\min} q^{i-1}$) wyraża się wzorem:

$$\begin{aligned} P(m \geq m_i) &= P(m_i \leq m < m_{i+1}) + P(m_{i+1} \leq m < m_{i+2}) + \dots = \\ &= \sum_{j=i}^{\infty} P_j = \frac{q_0^{-\alpha}}{1 + \frac{q_0^{-\alpha}}{1 - q^{-\alpha}}} q^{2\alpha} \sum_{j=i}^{\infty} (q^{-\alpha})^j = \\ &= \frac{q_0^{-\alpha} q^{-\alpha}}{1 - q^{-\alpha} + q_0^{-\alpha}} (q^{i-1})^{-\alpha} = \\ &= \frac{q_0^{-\alpha} q^{-\alpha}}{1 - q^{-\alpha} + q_0^{-\alpha}} \left(\frac{m_i}{m_{\min}} \right)^{-\alpha} = \Pi(m_i). \end{aligned} \quad (2.17)$$

Zgodnie z oczekiwaniami, otrzymaliśmy dopełnienie dystrybucyj wyrażające słabe prawo Pareto o wykładniku α . Wykładnik ten wyrażony jest poprzez wewnętrzną klasyfikację grupy gospodarstw domowych o wysokich dochodach jak też przez probabilistykę określającą szanse awansu lub degradacji wewnątrz tej grupy (patrz równanie (2.8)). Jak widać, model drabinowy określa warunki, w jakich budowany jest rozkład potęgowy względnych dochodów gospodarstw domowych o wysokich dochodach.

Średni dochód – podejście dynamiczne

Wróćmy teraz do równań (2.2) i (2.3) opisujących dynamikę liczby gospodarstw domowych, czyli ich liczbę $N_i(t)$ w chwili t w danej podgrupie i . W dalszym ciągu,

dokonajmy analizy pierwszego momentu, czyli średniego dochodu w chwili t :
 $\langle m(t) \rangle = \frac{1}{N} \sum_{i=1}^{\infty} N_i(t) m_i$.

Mnożąc stronami równanie (2.2) przez m_i oraz równanie (2.3) przez m_1 dostajemy:

$$m_i N_i(t) = m_{i+1} q^{-1} N_{i+1}(t-1) p_- + m_{i-1} q N_{i-1}(t-1) p_+, \quad i \geq 2, \quad (2.18)$$

$$m_1 N_1(t) = m_2 q^{-1} N_2(t-1) p_- + m_1 N_1(t-1) p_0, \quad (2.19)$$

gdzie wykorzystano tutaj fakt, że: $q^{i-1} q = q^i$ a zatem: $m_i q = m_{i+1}$. Dodając stronami równania (2.18) i (2.19), możemy wyrazić średni dochód w postaci:

$$\langle m(t) \rangle = p_- q^{-1} \langle m(t-1) \rangle + p_+ q \langle m(t-1) \rangle - \frac{p_- q^{-1}}{N} m_1 N_1(t-1) + \frac{p_0}{N} m_1 N_1(t-1). \quad (2.20)$$

Pamiętając, że $m_1 \equiv m_{\min}$ otrzymujemy:

$$\langle m(t) \rangle = (p_- q^{-1} + p_+ q) \langle m(t-1) \rangle + m_{\min} N_1(t-1) \frac{p_0 - p_- q^{-1}}{N}. \quad (2.21)$$

Zajmiemy się teraz analizą wartości średniego dochodu w chwili t , $\langle m(t) \rangle$, dla przypadku gdy $\alpha = 1$, gdyż, jak zostanie pokazane w rozdziale 3.3 oraz 4.4, gospodarstwa domowe o wysokich dochodach opisywane są (w przybliżeniu) prawem Pareto o takim właśnie wykładniku².

Gdy $\alpha = 1$, wówczas $p_- q^{-1} + p_+ q = 1$ (gdyż $p_- q^{-1} = p_+$ a $p_+ q = p_-$). Jeśli dodatkowo założymy, że $p_0 = p_- q^{-1}$, to wtedy średni dochód gospodarstw domowych o wysokich dochodach jest stały w czasie:

$$\langle m(t) \rangle = \langle m(t-1) \rangle. \quad (2.22)$$

Natomiast, gdy $p_0 - p_- q^{-1} > 0 \equiv \frac{p_0}{p_-} > q^{-1} \equiv p_0 > p_+$, to wówczas następuje wzrost średniego dochodu. W przeciwnym wypadku, dla $p_0 - p_- q^{-1} < 0$ mamy do czynienia z jego spadkiem.

Zauważmy, że jeśli prawdopodobieństwo przetrwania p_0 w podgrupie najniższej ($i = 1$) wzrasta (przy założeniu, że $p_0 > p_+$) to prawdopodobieństwo wyjścia z tej

² Oczywiście, nic nie stoi na przeszkodzie, aby rozważać sytuację ogólną ($\alpha \neq 1$), chociaż przypadek $\alpha = 1$ jest szczególnie interesujący.

podgrupy do podgrupy wyższej ($i = 2$) maleje. Wzrost p_0 pociąga za sobą wzrost N_1 , gdyż:

$$N_1 = N \left(1 + \frac{q_0^{-1}}{1 - q^{-1}} \right)^{-1} = N \frac{p_- - p_+}{p_- - p_+ + 1 - p_0} = N \frac{1 - \frac{p_+}{p_-}}{1 - \frac{p_+}{p_-} + \frac{1 - p_0}{p_-}} \quad (2.23)$$

oraz spadek N_i , ponieważ:

$$N_i = (q^{-\alpha})^{i-2} q_0^{-\alpha} N_1 = \left(\frac{p_+}{p_-} \right)^{i-2} \frac{1 - p_0}{p_-}, \quad i \geq 2. \quad (2.24)$$

Zatem, jak należało oczekiwać, wzrost p_0 ($p_0 > p_+$) powoduje zwiększenie liczby gospodarstw domowych w podgrupie $i = 1$ oraz spadek ich liczby w pozostałych podgrupach, pociągając jednocześnie za sobą (paradoksalnie) wzrost średniego dochodu dla całej klasy gospodarstw domowych o wysokich dochodach.

2.2 ROZSZERZONY MODEL YAKOVENKO

Wprowadzony w rozdziale 1.5 model Yakovenko jest w stanie opisać tylko niskie oraz średnie dochody gospodarstw domowych w Unii Europejskiej. Nie charakteryzuje on niestety wielce istotnych, wysokich dochodów gospodarstw domowych.

Zasadniczym celem niniejszej rozprawy doktorskiej jest sformułowanie modelu, dla którego dopełnienie dystrybuanty opisywałoby dochody uzyskiwane przez wszystkie trzy klasy społeczne tj. gospodarstwa domowe o niskich, średnich i wysokich dochodach, włączając w to obszary przejściowe pomiędzy tymi klasami.

Członkowie gospodarstw domowych o wysokich dochodach są bardzo często właścicielami firm bądź korporacji. Zatem, oczekujemy, że ich dochody będą opisywane (z dobrym przybliżeniem) analogicznie jak dochody firm, czyli prawem Zipfa³ [88,91,92].

³ Prawo Zipfa jest w istocie słabym prawem Pareto z wykładnikiem $\alpha = 1$. Model mikroskopowy (drabinowy) odtwarzający słabe prawo Pareto z wykładnikiem $\alpha = 1$ został omówiony w rozdziale 2.1.

W celu uzyskania gęstości prawdopodobieństwa opisującej dochody wszystkich trzech klas społecznych współczynniki $A(m)$ i $B(m)$ występujące w równaniu (1.10) przyjmujemy w następującej postaci:

$$A(m) = \begin{cases} A^{<}(m) = A_0 + a m, & \text{jeśli } m < m_1, \\ A^{\geq}(m) = A'_0 + a' m, & \text{jeśli } m \geq m_1, \end{cases} \quad (2.25)$$

$$B(m) = \begin{cases} B^{<}(m) = B_0 + b m^2 = b (m_0^2 + m^2) & \text{jeśli } m < m_1, \\ B^{\geq}(m) = B'_0 + b' m^2 = b' (m_0'^2 + m^2) & \text{jeśli } m \geq m_1, \end{cases} \quad (2.26)$$

gdzie⁴ $m_0^2 = B_0/b$ oraz $m_0'^2 = B'_0/b'$. Wprowadzony tutaj parametr m_1 traktujemy jako graniczną wartość dochodu pomiędzy gospodarstwami domowymi o średnich i wysokich dochodach (parametry m_0 oraz m_0' omawiamy poniżej). Zarówno wartość m_0 jak i m_1 wyznacza się na podstawie empirycznego dopełnienia dystrybucyj (patrz rozdział 4.4).

Przyjęcie współczynników $A(m)$ oraz $B(m)$, odpowiednio, w postaci (2.25) i (2.26) zakłada, że mechanizm zmian dochodu jest formalnie taki sam dla całego społeczeństwa, natomiast różnice pochodzą z innych proporcji pomiędzy procesem addytywnym a multiplikatywnym. Innymi słowy, mechanizm ten wyrażony poprzez progowe, nieliniowe równanie Langevina zdefiniowane wzorami⁵ (1.2), (2.25) oraz (2.26) może prowadzić do różnej dynamiki zmian dochodu – gospodarstwa domowe o niskich i średnich dochodach bogacą się (bądź tracą dochody) inaczej niż gospodarstwa domowe o wysokich dochodach, ponieważ współczynniki $A^{<}(m)$, $A^{\geq}(m)$ oraz $B^{<}(m)$, $B^{\geq}(m)$ różnią się od siebie znacząco (w ogólności $A_0 \neq A'_0$, $a \neq a'$, $B_0 \neq B'_0$, $b \neq b'$).

Aby gęstość prawdopodobieństwa nie miała nieciągłości w punkcie $m = m_1$ musi zachodzić⁶:

$$P^{<}(m = m_1) = P^{\geq}(m = m_1), \quad (2.27)$$

⁴ Pojawiającego się tutaj parametru m_0 nie należy mylić z parametrem m_0 występującym w prawie Pareto. Przypomnijmy, że w przypadku silnego prawa Pareto parametr m_0 oznacza graniczną, najniższą wartość dochodu, jaką mogą osiągać gospodarstwa domowe, natomiast w słabym prawie Pareto jest on parametrem swobodnym i nie posiada interpretacji.

⁵ Przypomnijmy, że: $B(m) = \frac{C^2(m)}{2}$.

⁶ Przyjęcie warunku, że rozkład gęstości prawdopodobieństwa dochodów uzyskiwanych przez gospodarstwa domowe jest ciągły wynika z zasady zwanej brzytwą Ockhama [93].

gdzie na podstawie (1.35) możemy zapisać, że⁷:

$$P^{<}(m) = \frac{\text{const}}{B^{<}(m)} \exp \left(- \int_0^{m^{(<m_1)}} \frac{A^{<}(m')}{B^{<}(m')} dm' \right) \quad (2.28)$$

oraz

$$\begin{aligned} P^{\geq}(m) &= \frac{\text{const}}{B^{\geq}(m)} \exp \left(- \int_0^{m^{(\geq m_1)}} \frac{A(m')}{B(m')} dm' \right) = \\ &= \frac{\text{const}}{B^{\geq}(m)} \exp \left(- \int_0^{m_1} \frac{A^{<}(m')}{B^{<}(m')} dm' \right) \times \exp \left(- \int_{m_1}^{m^{(\geq m_1)}} \frac{A^{\geq}(m')}{B^{\geq}(m')} dm' \right). \end{aligned} \quad (2.29)$$

Podstawiając równanie (2.28) oraz (2.29) do równania (2.27) otrzymujemy:

$$\begin{aligned} &\frac{\text{const}}{B^{<}(m = m_1)} \exp \left(- \int_0^{m=m_1} \frac{A^{<}(m')}{B^{<}(m')} dm' \right) = \\ &= \frac{\text{const}}{B^{\geq}(m = m_1)} \exp \left(- \int_0^{m_1} \frac{A^{<}(m')}{B^{<}(m')} dm' \right) \times \exp \left(- \int_{m_1}^{m_1} \frac{A^{\geq}(m')}{B^{\geq}(m')} dm' \right), \end{aligned} \quad (2.30)$$

a stąd:

$$B^{<}(m = m_1) = B^{\geq}(m = m_1) \Leftrightarrow B_0 + b m_1^2 = B'_0 + b' m_1^2. \quad (2.31)$$

Ponieważ m_0 i m'_0 są wartościami dochodu wyznaczającymi granicę pomiędzy gospodarstwami domowymi o niskich i średnich dochodach, więc musimy przyjąć, że:

$$m'_0 = m_0 \Leftrightarrow \frac{B_0}{b} = \frac{B'_0}{b'}. \quad (2.32)$$

Upraszczając dodatkowo nasze rozważania zakładamy, że:

$$B'_0 = B_0 \text{ and } b' = b, \quad (2.33)$$

co prowadzi do amplitudy szumu gaussowskiego $C(m)$ identycznej dla wszystkich gospodarstw domowych. Zatem gospodarstwa domowe o niskich i średnich dochodach na poziomie dynamiki zmian dochodów różnią się od gospodarstw domowych o wysokich dochodach tylko współczynnikiem dryfu $A(m)$. Jest to ważny wniosek wskazujący na główny mechanizm różnicujący dochody klas społecznych.

⁷ Podobnie jak poprzednio za dolną granicę całkowania we wzorze (1.35) wybieramy zero, ponieważ w praktyce dla danych empirycznych m_p jest bardzo małe, rzędu 1 euro lub 1 złotówki.

Uwzględniając powyższe założenia możemy napisać, że:

$$B(m) = B_0 + bm^2 = b(m_0^2 + m^2). \quad (2.34)$$

Podstawiając (2.25) oraz (2.34) do (1.35) otrzymujemy⁸ dla $m < m_1$:

$$\begin{aligned} P(m) &= \frac{\text{const}}{b(m_0^2 + m^2)} \exp\left(-\int_0^m \frac{A_0 + am'}{b(m_0^2 + m'^2)} dm'\right) = \\ &= C' \frac{\exp\left(-\frac{m_0}{T} \arctan \frac{m}{m_0}\right)}{\left[1 + \left(\frac{m}{m_0}\right)^2\right]^{\frac{\alpha+1}{2}}}, \end{aligned} \quad (2.35)$$

gdzie $T = \frac{B_0}{A_0}$, $\frac{\alpha+1}{2} = \frac{a}{2b} + 1$ oraz $C' = \frac{\text{const}}{bm_0^2}$. Z kolei, dla $m \geq m_1$ mamy:

$$\begin{aligned} P(m) &= \frac{\text{const}}{b(m_0^2 + m^2)} \exp\left(-\int_0^{m_1} \frac{A_0 + am'}{b(m_0^2 + m'^2)} dm'\right) \times \exp\left(-\int_{m_1}^m \frac{A'_0 + a'm'}{b(m_0^2 + m'^2)} dm'\right) = \\ &= C'' \frac{\exp\left(-\frac{m_0}{T_1} \arctan \frac{m}{m_0}\right)}{\left[1 + \left(\frac{m}{m_0}\right)^2\right]^{\frac{\alpha_1+1}{2}}}, \end{aligned} \quad (2.36)$$

gdzie $T_1 = \frac{B'_0}{A'_0}$, $\frac{\alpha_1+1}{2} = \frac{a'}{2b} + 1$ oraz

$$\begin{aligned} C'' &= \frac{\text{const}}{bm_0^{2+a'/b}} \exp\left(-\int_0^{m_1} \frac{A_0 + am'}{b(m_0^2 + m'^2)} dm'\right) \times \exp\left(\frac{A'_0}{bm_0} \arctan \frac{m_1}{m_0} + \frac{a'}{2b} \ln(m_0^2 + m_1^2)\right) = \\ &= C' \exp\left(-\left(\frac{1}{T} - \frac{1}{T_1}\right) m_0 \arctan \frac{m_1}{m_0}\right) \left[1 + \left(\frac{m_1}{m_0}\right)^2\right]^{(\alpha_1 - \alpha)/2}, \text{ przy czym } C' \text{ zostało zdefiniowane we wzorze (2.35)}. \end{aligned}$$

Podsumowując, gęstość prawdopodobieństwa, $P(m)$, opisująca dochody uzyskiwane przez wszystkie trzy klasy społeczne w Unii Europejskiej, tzn. gospodarstwa domowe o niskich, średnich i wysokich dochodach, przybiera postać:

$$P(m) = \begin{cases} C' \frac{\exp\left(-\frac{m_0}{T} \arctan \frac{m}{m_0}\right)}{\left[1 + \left(\frac{m}{m_0}\right)^2\right]^{\frac{\alpha+1}{2}}}, & \text{dla } m < m_1, \\ C'' \frac{\exp\left(-\frac{m_0}{T_1} \arctan \frac{m}{m_0}\right)}{\left[1 + \left(\frac{m}{m_0}\right)^2\right]^{\frac{\alpha_1+1}{2}}}, & \text{dla } m \geq m_1. \end{cases} \quad (2.37)$$

⁸ Podobnie jak poprzednio za dolną granicę całkowania we wzorze (1.35) wybieramy zero, ponieważ w praktyce m_p jest bardzo małe, rzędu 1 euro lub 1 złotówki.

W powyższym rozkładzie liczba efektywnych parametrów jest w istocie zredukowana, gdyż występują one w postaci ilorazów parametrów definiujących dynamikę zmian dochodów opisywaną równaniem Langevina (1.2). **Formuła (2.37) jest jednym z kluczowych wyników teoretycznych niniejszej rozprawy doktorskiej.**

Dla $m \ll m_0$ rozkład (2.37) (a dokładniej, pierwsze wyrażenie) przechodzi w prawo Boltzmann–Gibbsa. Natomiast, dla $m_1 \gg m_0$ przechodzi on kolejno (dla $m_0 \ll m < m_1$) w słabe prawo Pareto z wykładnikiem α , a następnie (dla $m \geq m_1$) w słabe prawo Pareto z wykładnikiem α_1 .

Podobnie jak w przypadku modelu Yakovenko, także dla rozszerzonego modelu Yakovenko wyrażonego przez wzór (2.37) nie daje się wyznaczyć dopełnienia dystrybucyjnego $\Pi(m)$ w postaci analitycznej. Efektywna metoda numeryczna została przedstawiona w rozdziale 4.4.

Należy podkreślić, że rozszerzony model Yakovenko dotyczy Unii Europejskiej jako całości (patrz rozdział 4.4, rysunki 4.12 – 4.13). Do opisu polskich gospodarstw domowych konieczne jest inne podejście.

ANALIZA ROCZNYCH DOCHODÓW GOSPODARSTW DOMOWYCH W POLSCE

*„To, co nazywamy fizyką,
obejmuje całą grupę nauk przyrodniczych,
które opierają swe teorie na pomiarach
i których idee i twierdzenia
dają się sformułować za pomocą matematyki.”*

ALBERT EINSTEIN

3.1 OPIS BAZY DANYCH

W analizie danych empirycznych wykorzystano dane Głównego Urzędu Statystycznego¹ dotyczące rocznych² dochodów rozporządzalnych gospodarstw domowych w Polsce w latach 2004–2010. Po usunięciu rekordów (czyli tych dochodów w liczbie 264 w 2004 r., 198 w 2005 r., 226 w 2006 r., 245 w 2007 r., 251 w 2008 r., 271 w 2009 r., 223 w 2010 r.), dla których dochód był mniejszy (czyli strata) bądź równy zero, baza danych zawierała odpowiednio 32054, 34569, 37282, 37121, 37107, 37031 i 37189 rekordów. Niedodatnie wartości dochodu uznano za obserwacje błędne, gdyż zgodnie z definicją dochód rozporządzalny to (cytuje za [94]) „... suma bieżących dochodów gospodarstw domowych z poszczególnych źródeł

¹ <http://www.stat.gov.pl/>

² Główny Urząd Statystyczny gromadzi dane o dochodach miesięcznych gospodarstw domowych. Aby uzyskać dochody roczne, dochody miesięczne przemnożono przez czynnik 12.

pomniejszona o zaliczki na podatek dochodowy od osób fizycznych płacone przez płatnika w imieniu podatnika (od dochodów z pracy najemnej oraz od niektórych świadczeń z ubezpieczenia społecznego i świadczeń pozostałych), o podatki od dochodów z własności, podatki płacone przez osoby pracujące na własny rachunek, w tym przedstawicieli wolnych zawodów i osób użytkujących gospodarstwo indywidualne w rolnictwie oraz o składki na ubezpieczenie społeczne i zdrowotne. W skład dochodu rozporządzalnego wchodzi dochody pieniężne i niepieniężne, w tym spożycie naturalne (towary lub usługi konsumpcyjne pobrane na potrzeby gospodarstwa domowego z gospodarstwa indywidualnego bądź z prowadzonej działalności gospodarczej na własny rachunek – rolniczej i pozarolniczej) oraz towary i usługi otrzymane bezpłatnie. Dochód rozporządzalny jest przeznaczony na wydatki oraz przyrost oszczędności”.

3.2 EMPIRYCZNE DOPEŁNIENIA DYSTRYBUANT

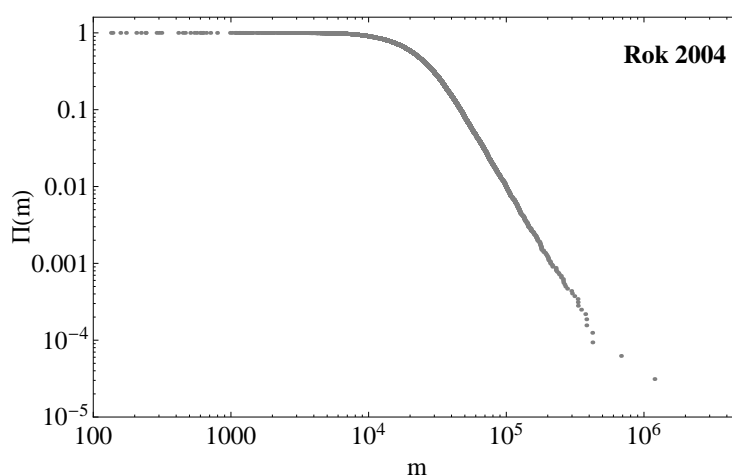
Jednym z celów niniejszej rozprawy doktorskiej jest analiza empirycznych dopełnień dystrybuant rocznych dochodów rozporządzalnych gospodarstw domowych w Polsce. Rozkłady te skonstruowano zgodnie z następującymi regułami:

- Po pierwsze dane o dochodach zostały uszeregowane w rankingu, czyli malejąco – od gospodarstwa domowego o największym dochodzie do gospodarstwa domowego o najmniejszym dochodzie.
- Po drugie, dla każdego gospodarstwa policzono, zgodnie z formułą Weibulla³ [95–97], ułamek: $\frac{l}{n+1}$, gdzie l jest numerem miejsca (czyli rangą) jakie zajmuje dane gospodarstwo domowe w tym rankingu, a n jest liczbą gospodarstw domowych w analizowanej próbie. Otrzymane wartości oznaczają ułamek gospodarstw domowych uzyskujących dochód równy, bądź większy niż gospodarstwo domowe znajdujące się w rankingu na pozycji l .
- Następnie sporządzono wykres, na którym na osi odciętych odłożono dochód gospodarstwa domowego a na osi rzędnych policzony dla niego (powyżej zdefiniowany) ułamek.

³ Metoda Weibulla została dokładniej opisana w dodatku D.

Dopełnienie dystrybuanty otrzymane w ten sposób jest stabilne⁴, tzn. obarczone wyraźnie mniejszym błędem⁵ niż sama gęstość prawdopodobieństwa, przy czym jednocześnie nadaje się ono do analizy znaczących zmian w czasie struktury dochodów gospodarstw domowych. Rozkład taki zawiera tyle samo punktów empirycznych (gospodarstw domowych), co pierwotne dane. Dzięki temu, kolejne analizy mogą być wykonywane na próbach o maksymalnie dostępnej liczebności.

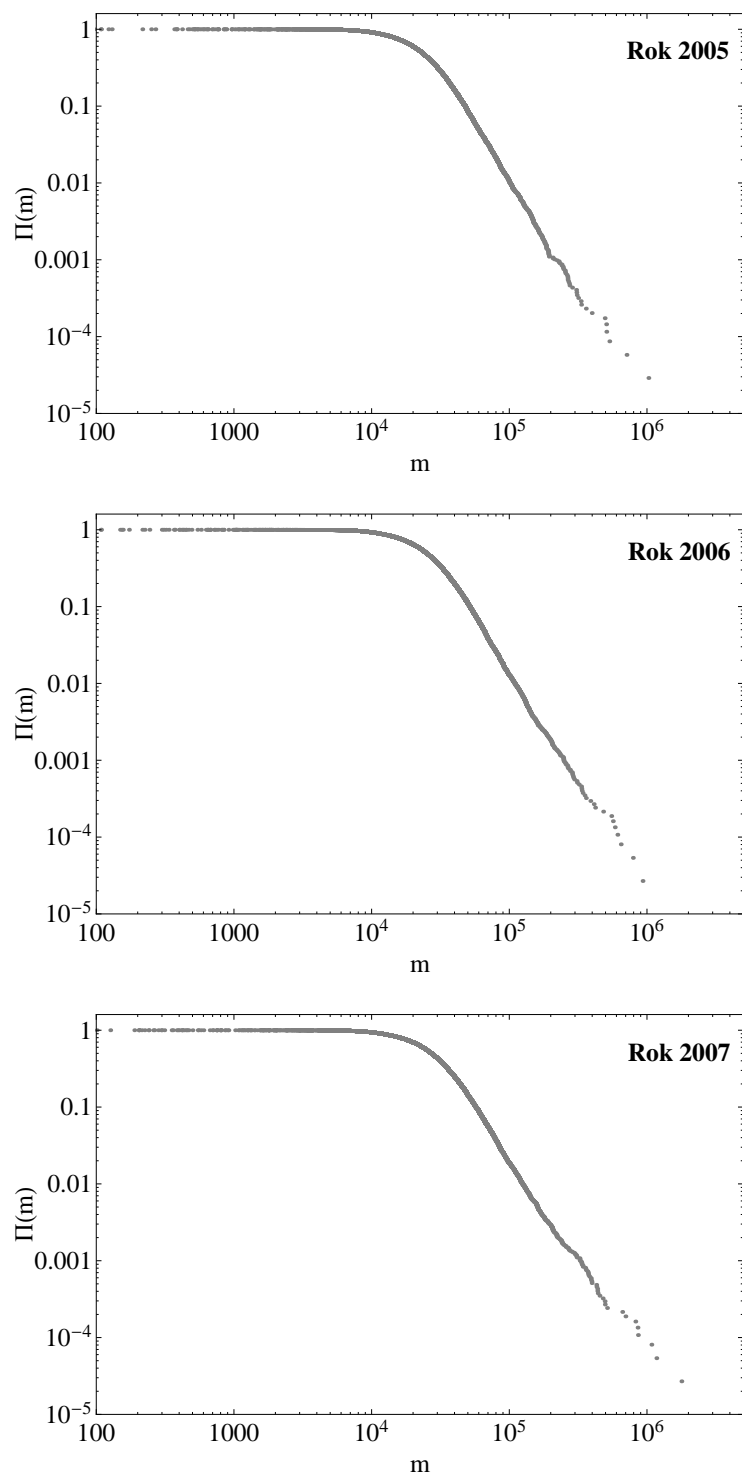
Empiryczne dopełnienia dystrybuant rocznych dochodów rozporządzalnych gospodarstw domowych w Polsce zostały przedstawione na rysunkach 3.1 – 3.3 (skala logarytmiczna).



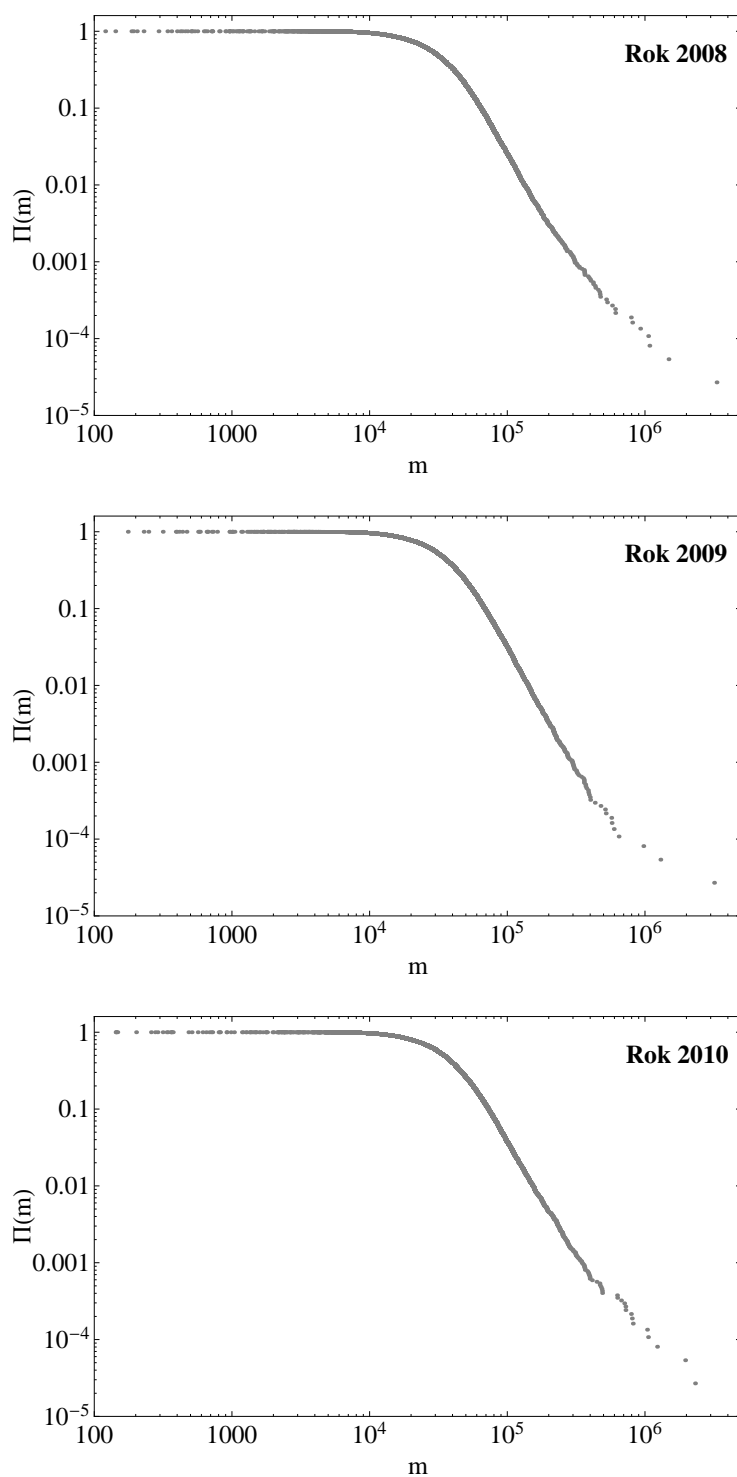
Rysunek 3.1: Empiryczne dopełnienie dystrybuanty, $\Pi(m)$, rocznych dochodów rozporządzalnych gospodarstw domowych w Polsce, m , z roku 2004. Właśnie takie rozkłady zostaną w dalszej części pracy poddane analizie teoretycznej. Cechują się one dużą stabilnością w czasie.

⁴ Nie mylić z definicją rozkładu stabilnego, np. Gaussa lub Lévy'ego.

⁵ W przypadku np. gdy dwa gospodarstwa domowe uzyskują ten sam dochód, to w rankingu otrzymują różne miejsca, a co za tym idzie inne są też wartości obliczanego ułamka. Fakt, ten ma znaczenie marginalne i nie wpływa na kształt otrzymywanych empirycznych dopełnień dystrybuant. Należy pamiętać, że dane o dochodach mają charakter szacunkowy i te same wartości dochodu można sztucznie zróżnicować np. o jeden grosz.



Rysunek 3.2: Empiryczne dopełnienia dystrybuant, $\Pi(m)$, rocznych dochodów rozporządzalnych gospodarstw domowych w Polsce, m , z lat 2005–2007. Także te rozkłady zostaną w dalszej części pracy poddane analizie teoretycznej. Cechują się one dużą stabilnością w czasie (podobnie jak ten z roku 2004).



Rysunek 3.3: Empiryczne dopełnienia dystrybuant, $\Pi(m)$, rocznych dochodów rozporządzalnych gospodarstw domowych w Polsce, m , z lat 2008–2010. Również te rozkłady zostaną w dalszej części pracy poddane analizie teoretycznej. Cechują się one (tak jak poprzednie z lat 2004–2007) dużą stabilnością w czasie.

3.3 ANALIZA DANYCH EMPIRYCZNYCH

W kolejnych rozdziałach porównuję empiryczne dopełnienia dystrybuant, dotyczące rocznych dochodów rozporządzalnych gospodarstw domowych w Polsce, z przewidywaniami modeli teoretycznych, tj. z Prawem Efektów Proporcjonalnych, słabym prawem Pareto oraz uogólnionym modelem Lotka–Volterra. Następnie, omawiam metodę wyznaczania wartości wykładnika Pareto (pochodzącego ze słabego prawa Pareto) przy pomocy analizy rankingu dochodów. Umożliwia to nie tylko wyznaczenie niezależną metodą wykładnika Pareto dla gospodarstw domowych uzyskujących średnie dochody, ale pozwala również określić jego wartość dla gospodarstw domowych osiągających wysokie dochody. Wyznaczenie wykładnika Pareto dla gospodarstw domowych osiągających wysokie dochody na podstawie empirycznych dopełnień dystrybuant nie jest możliwe, gdyż rozkłady te zawierają tylko kilka punktów empirycznych reprezentujących tego typu gospodarstwa.

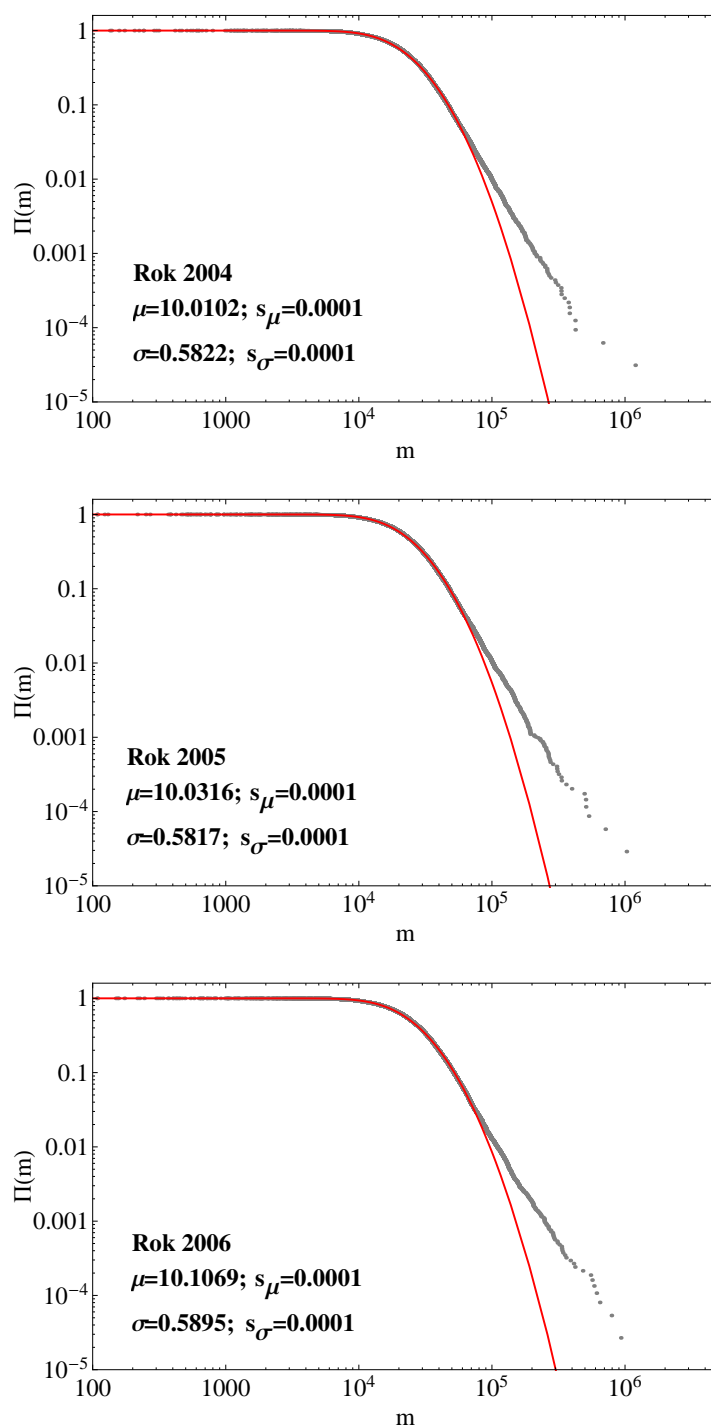
Wykorzystanie Prawa Efektów Proporcjonalnych

W pierwszej kolejności do empirycznych dopełnień dystrybuant dotyczących rocznych dochodów rozporządzalnych gospodarstw domowych w Polsce dopasowano dopełnienie dystrybuanty rozkładu log-normalnego (1.76) wynikające z Prawa Efektów Proporcjonalnych (patrz rozdział 1.6):

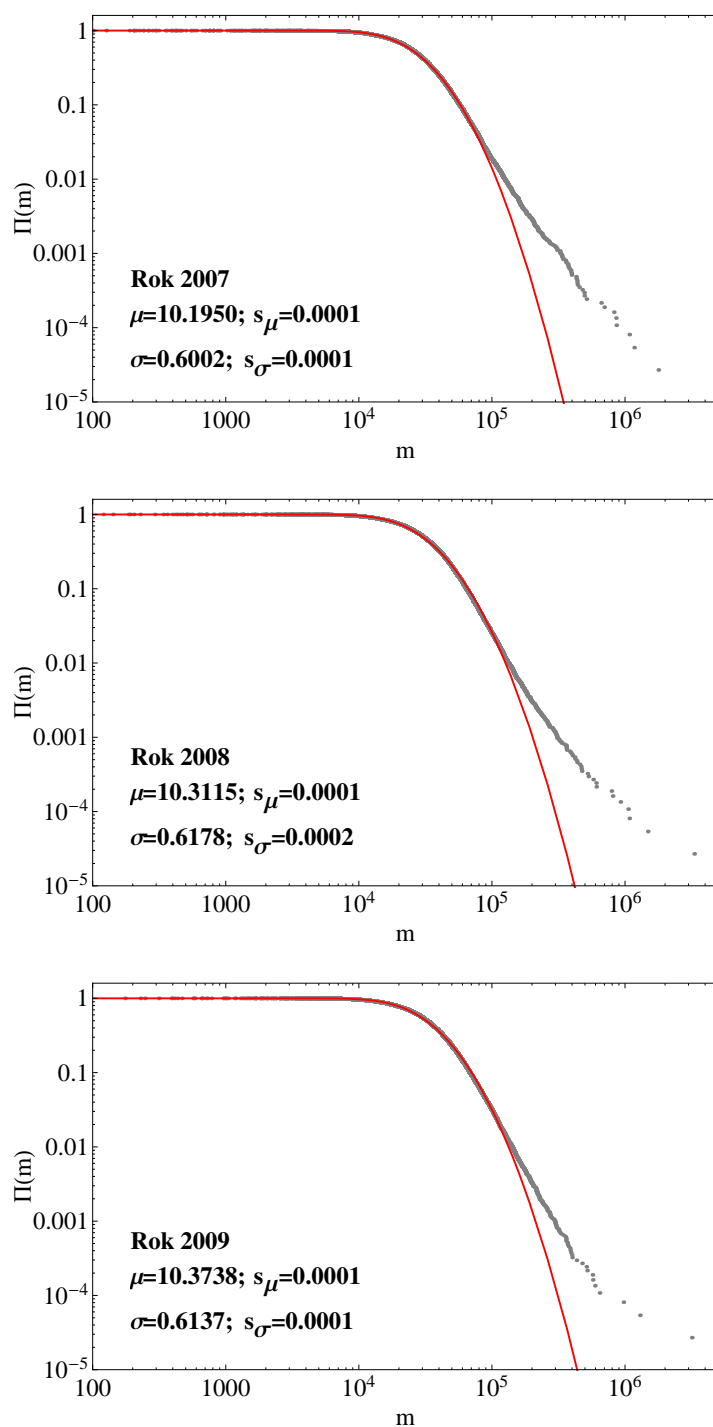
$$\Pi(m) = \frac{1}{2} \left[1 - \operatorname{erf} \left(\frac{\ln m - \mu}{\sqrt{2}\sigma} \right) \right]. \quad (3.1)$$

Wyniki w skali logarytmicznej zostały przedstawione na rysunkach 3.4 – 3.6⁶.

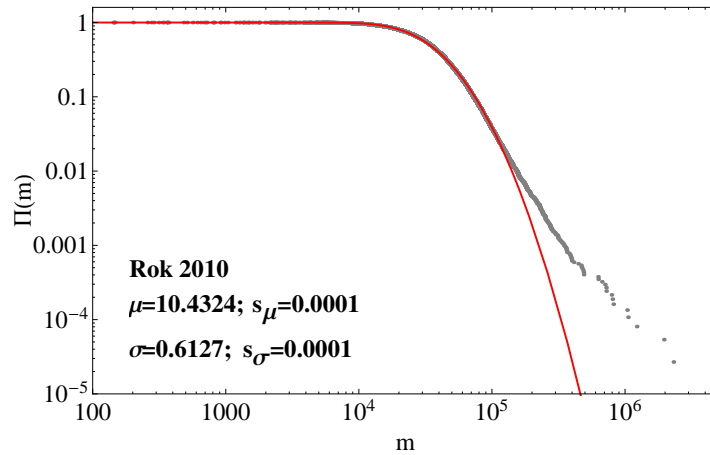
6 Występujące tutaj i na dalszych rysunkach oznaczenie s_x jest odchyleniem standardowym parametru x . Dopełnienie dystrybuanty rozkładu log-normalnego dopasowano do wszystkich danych empirycznych dotyczących dochodów gospodarstw domowych w Polsce. W celu dopasowania krzywych posłużono się algorytmem Levenberga–Marquardta [98].



Rysunek 3.4: Dopasowanie dopełnienia dystrybuanty rozkładu log-normalnego (linia ciągła) do empirycznego dopełnienia dystrybuanty rocznych dochodów rozporządzalnych gospodarstw domowych w Polsce (punkty) z lat 2004–2006. Jak widać, Prawo Efektów Proporcjonalnych całkiem dobrze opisuje dane empiryczne dla niskich dochodów gospodarstw domowych.



Rysunek 3.5: Dopasowanie dopełnienia dystrybuanty rozkładu log-normalnego (linia ciągła) do empirycznego dopełnienia dystrybuanty rocznych dochodów rozporządzalnych gospodarstw domowych w Polsce (punkty) z lat 2007–2009. Jak widać, Prawo Efektów Proporcjonalnych całkiem dobrze opisuje dane empiryczne dla niskich dochodów gospodarstw domowych.



Rysunek 3.6: Dopasowanie dopełnienia dystrybuanty rozkładu log-normalnego (linia ciągła) do empirycznego dopełnienia dystrybuanty rocznych dochodów rozporządzalnych gospodarstw domowych w Polsce (punkty) z roku 2010. Jak widać, Prawo Efektów Proporcjonalnych całkiem dobrze opisuje dane empiryczne dla niskich dochodów gospodarstw domowych.

Jak wynika z wykresów zamieszczonych na rysunkach 3.4 – 3.6, dopełnienie dystrybuanty rozkładu log-normalnego całkiem dobrze opisuje gospodarstwa domowe o niskich dochodach. Wartości parametrów dla poszczególnych dopasowań zostały zebrane w tabeli 3.1⁷.

⁷ Błędy parametrów s_x , $x = \mu, \sigma$, zostały podane w legendach na rysunkach 3.4 – 3.6. Jak widać, są one bardzo małe. Z kolei, średnią $\langle x \rangle$ oraz rozpiętość R_x można oszacować na podstawie wartości wyznaczonych parametrów.

Tabela 3.1: Parametry uzyskane z dopasowania dopełnienia dystrybuanty rozkładu log-normalnego do empirycznego dopełnienia dystrybuanty rocznych dochodów rozporządzalnych gospodarstw domowych w Polsce z lat 2004–2010. W przedostatnim wierszu znajdują się wartości średnie parametrów. Możemy zauważyć, że rozrzut względny (względna rozpiętość) R_x parametrów nie przekracza 6%.

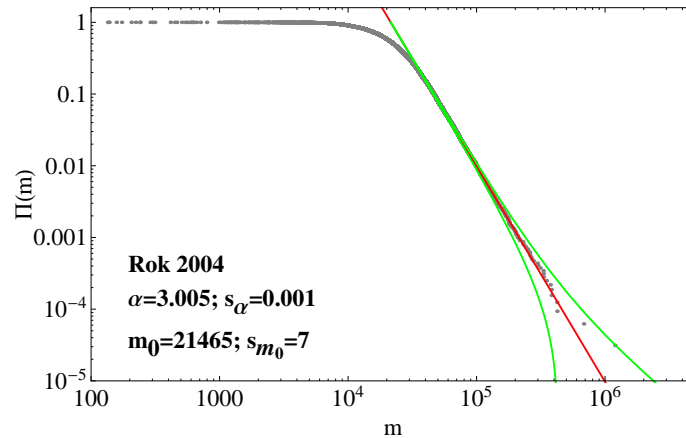
Rok	μ	σ	$\frac{\sigma}{\mu}$
2004	10.0102	0.5822	0.0582
2005	10.0316	0.5817	0.0580
2006	10.1069	0.5895	0.0583
2007	10.1950	0.6002	0.0589
2008	10.3115	0.6178	0.0599
2009	10.3738	0.6137	0.0592
2010	10.4324	0.6127	0.0587
$\langle x \rangle$	10.2088	0.5997	0.0587
$R_x = \frac{x_{\max} - x_{\min}}{\langle x \rangle}$	0.04	0.06	0.03

Jak widać, analiza klasy społecznej reprezentowanej przez gospodarstwa domowe o niskich dochodach nie dostarcza informacji o kryzysie – klasa ta jest niezwykle stabilna, gdyż zarówno parametry μ jak i σ są wolnozmiennymi funkcjami czasu.

Wykorzystanie słabego prawa Pareto

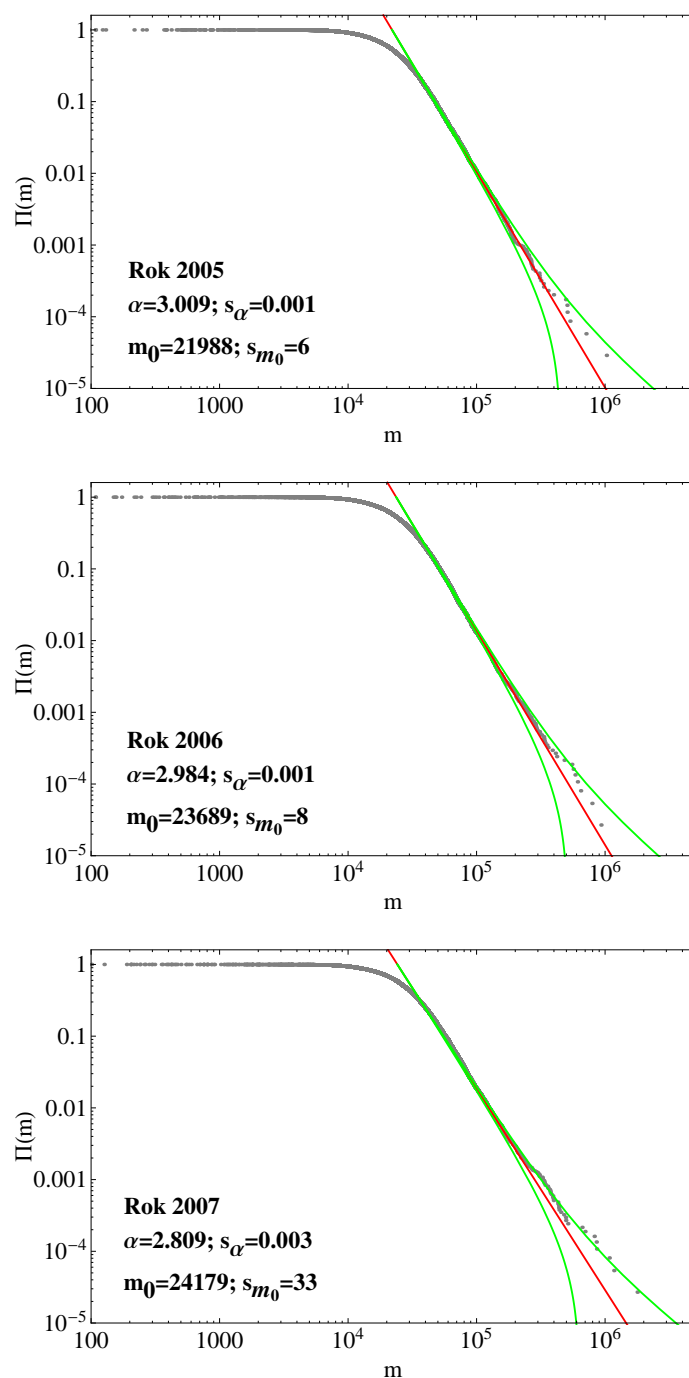
Do „ogonów” omawianych empirycznych dopełnień dystrybuant dopasowano (patrz rys. 3.7 – 3.9) również słabe prawo Pareto⁸ (przedstawione w rozdziale 1.4) postaci (przypomnijmy wzór (1.61)):

$$\Pi(m) \sim (m/m_0)^{-\alpha}. \quad (3.2)$$

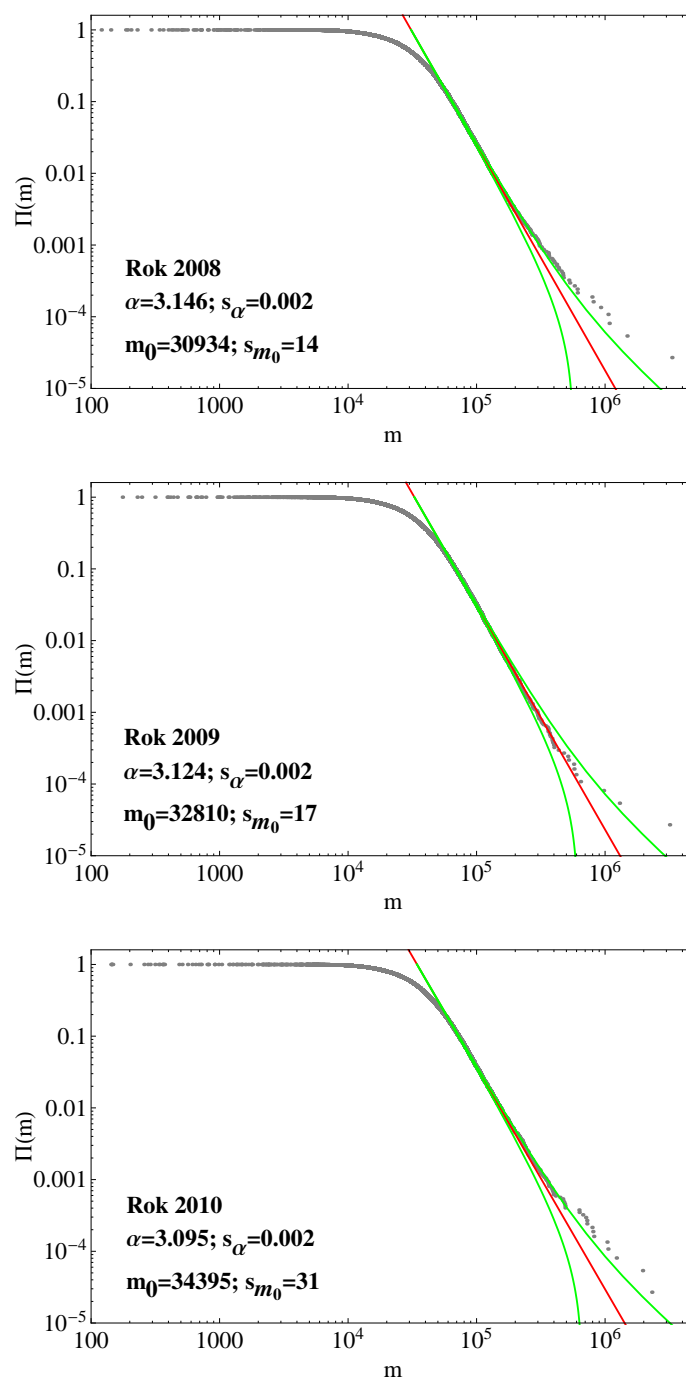


Rysunek 3.7: Dopasowanie słabego prawa Pareto (linia ciągła) do empirycznego dopełnienia dystrybuanty rocznych dochodów rozporządzalnych gospodarstw domowych w Polsce (punkty) z roku 2004. Słabe prawo Pareto dobrze opisuje gospodarstwa domowe o średnich dochodach. Dyskusja dotycząca zbioru punktów układających się powyżej słabego prawa Pareto została przeprowadzona w dalszej części pracy. Zielone linie oznaczają przedziały ufności dla zebranych danych empirycznych na poziomie 95%.

⁸ W celu dopasowania krzywych posłużono się algorytmem Levenberga–Marquardta [98]. Dopełnienie dystrybuanty słabego prawa Pareto dopasowano do danych empirycznych dotyczących średnich dochodów gospodarstw domowych w Polsce. Wyboru podpróbki dokonano tak, aby do ogonów analizowanych rozkładów można było dopasować prostą w skali log-log. W tym celu należało znaleźć graniczny dochód, poniżej którego dane są odrzucane. Zatem dopasowanie przeprowadzono w zakresie danych – od pewnego granicznego dochodu do dochodu najbogatszego gospodarstwa domowego dostępnego w próbie. Przy wyborze podpróbki, do której fitowano prostą (w skali log-log) kierowano się maksymalizacją wartości R^2 .



Rysunek 3.8: Dopasowanie słabego prawa Pareto (linia ciągła) do empirycznego dopełnienia dystrybucyjności rocznych dochodów rozporządzalnych gospodarstw domowych w Polsce (punkty) z lat 2005–2007. Słabe prawo Pareto dobrze opisuje gospodarstwa domowe o średnich dochodach. Dyskusja dotycząca zbioru punktów układających się powyżej słabego prawa Pareto została przeprowadzona w dalszej części pracy. Zielone linie oznaczają przedziały ufności dla zebranych danych empirycznych na poziomie 95%.



Rysunek 3.9: Dopasowanie słabego prawa Pareto (linia ciągła) do empirycznego dopełnienia dystrybucyjności rocznych dochodów rozporządzalnych gospodarstw domowych w Polsce (punkty) z lat 2008–2010. Słabe prawo Pareto dobrze opisuje gospodarstwa domowe o średnich dochodach. Dyskusja dotycząca zbioru punktów układających się powyżej słabego prawa Pareto została przeprowadzona w dalszej części pracy. Zielone linie oznaczają przedziały ufności dla zebranych danych empirycznych na poziomie 95%.

Słabe prawo Pareto dobrze opisuje gospodarstwa domowe o średnich dochodach. Na rysunkach 3.7 – 3.9 zaznaczono również przedziały ufności (zielone linie) dla zebranych danych empirycznych na poziomie 95%. Przedziały te wyznaczono przy użyciu metody przedstawionej w pracy [99]. Jak widać, niektóre z punktów empirycznego dopełnienia dystrybucyjności leżą poza przedziałami ufności. Zatem mają miejsce odstępstwa od słabego prawa Pareto. Jest to spowodowane obecnością w danych empirycznych gospodarstw domowych o wysokich dochodach. Ich liczba jest jednak zbyt mała, aby mogła zostać poddana opisowi statystycznemu. Gospodarstwa te zostaną przeanalizowane osobno, w dalszej części pracy przy użyciu rankingów.

Wartości wykładników Pareto dla poszczególnych lat zebrano w tabeli 3.2⁹.

Tabela 3.2: Wykładniki Pareto uzyskane z dopasowania słabego prawa Pareto do empirycznego dopełnienia dystrybucyjności rocznych dochodów rozporządzalnych gospodarstw domowych w Polsce z lat 2004–2010. W przedostatnim wierszu znajduje się wartość średnia wykładnika Pareto. Możemy zauważyć, że jego rozrzut względny (względna rozpiętość) R_x wynosi 11%.

Rok	α
2004	3.005
2005	3.009
2006	2.984
2007	2.809
2008	3.146
2009	3.124
2010	3.095
$\langle x \rangle$	3.025
$R_x = \frac{x_{\max} - x_{\min}}{\langle x \rangle}$	0.11

⁹ Błędy wykładników zostały zamieszczone w legendach na rysunkach 3.7 – 3.9. Z kolei, średnią $\langle x \rangle$ oraz rozpiętość R_x można oszacować na podstawie wartości wyznaczonych wykładników α .

Wykładnik Pareto może stanowić wskaźnik nierówności społecznych. Im mniejsza jest jego wartość, tym większe są nierówności uzyskiwanych dochodów w danej klasie społecznej. Z tabeli 3.2 wynika, że wykładnik Pareto osiągnął najmniejszą wartość w roku 2007, czyli roku „wybuchu” światowego kryzysu finansowego. **Zatem w roku tym rozwarstwienie społeczne osiągnęło wartość największą (najmniejszy wykładnik Pareto).**

Słabe prawo Pareto jest odpowiednie do zobrazowania gospodarstw domowych o średnich dochodach, ponieważ w tym wypadku zaczyna odgrywać rolę konkurencja rynkowa. Firmy konkurują między sobą a ich zyski przekładają się na zarobki pracowników – jednostek tworzących gospodarstwa domowe o średnich dochodach. W pracy [100] rozważono trzy typy konkurencji rynkowej. Następnie, przy pomocy symulacji komputerowych wykazano, że dochody najbogatszych (np. firm, gospodarstw domowych, jednostek, itp.), podlegających ustalonym zasadom konkurencji opisywane są prawem Pareto [100]. Zatem możemy oczekiwać, że także zarobki ich pracowników (jednostek tworzących gospodarstwa domowe o średnich dochodach) będą opisywane słabym prawem Pareto.

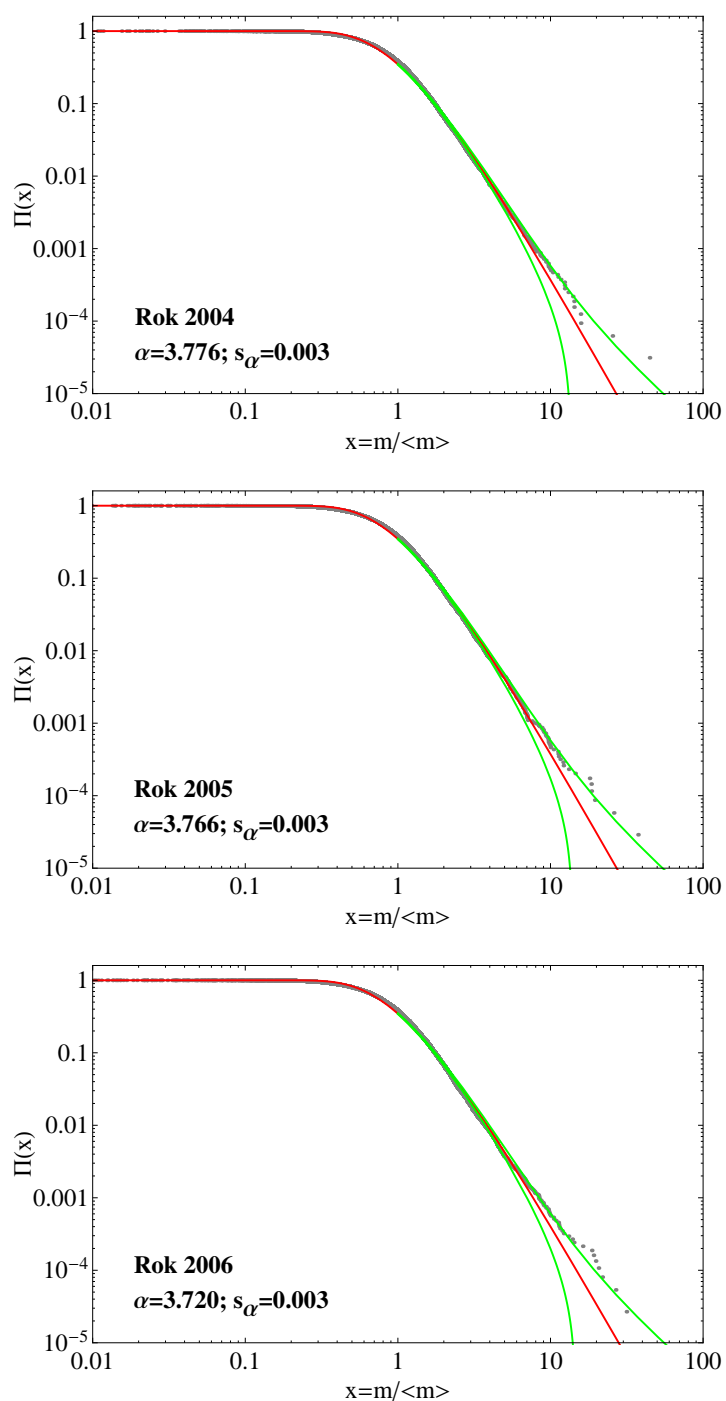
Wykorzystanie uogólnionego modelu Lotka–Volterra

W kolejnym kroku do rozważanych empirycznych dopełnień dystrybuant dopasowano dopełnienie dystrybuanty rozkładu (1.94) pochodzącego z uogólnionego modelu Lotka–Volterra (patrz rozdział 1.7):

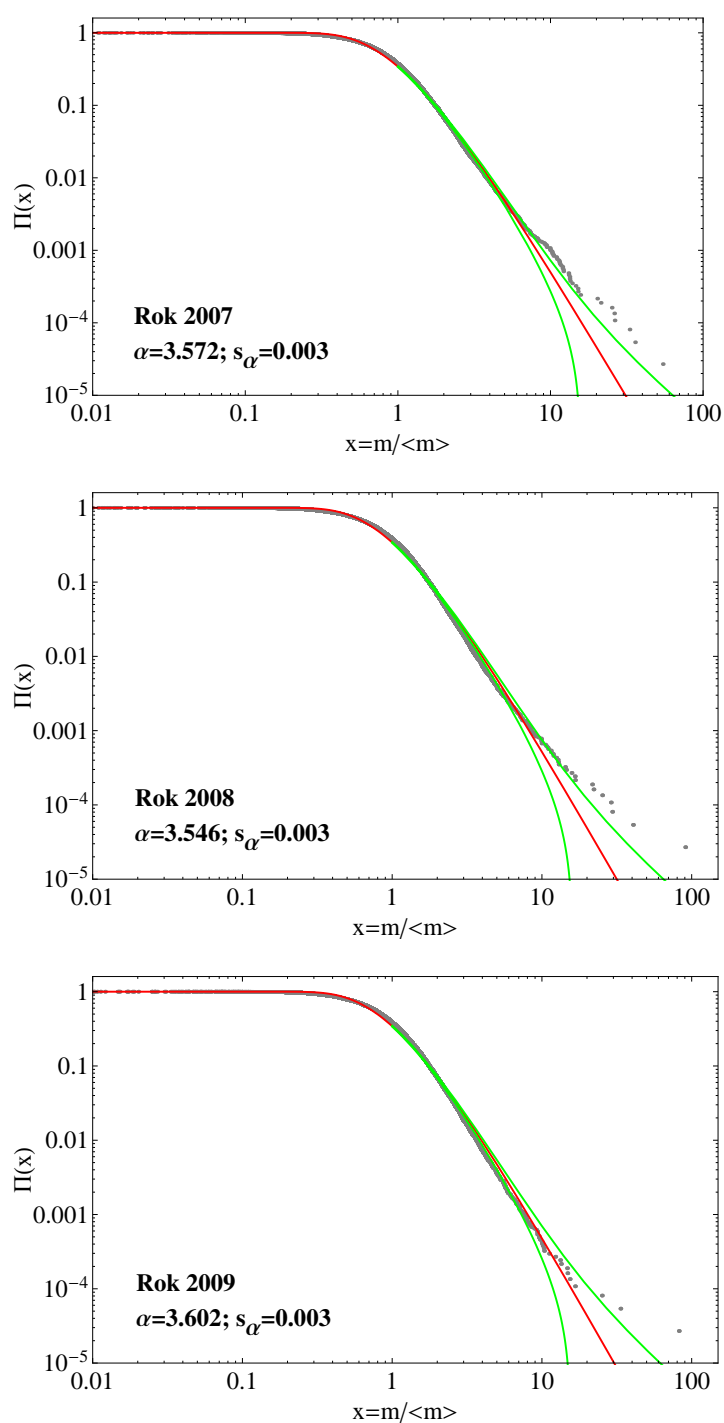
$$\Pi(x) = 1 - \frac{\Gamma(\alpha, \frac{\alpha-1}{x})}{\Gamma(\alpha)}. \quad (3.3)$$

Wyniki dopasowań znajdują się na rysunkach 3.10 – 3.12¹⁰ (skala log–log):

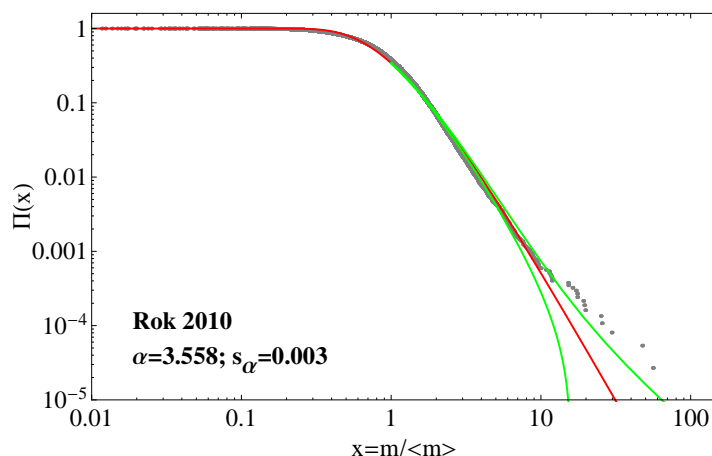
¹⁰ W celu dopasowania krzywych posłużono się algorytmem Levenberga–Marquardta [98]. Dopełnienie dystrybuanty uogólnionego modelu Lotka–Volterra dopasowano do wszystkich danych empirycznych dotyczących dochodów gospodarstw domowych w Polsce.



Rysunek 3.10: Dopasowanie uogólnionego modelu Lotka–Volterra (linia ciągła) do empirycznego dopełnienia dystrybuanty rocznych dochodów rozporządzalnych gospodarstw domowych w Polsce (punkty) z lat 2004–2006. Widać, że model ten dobrze opisuje gospodarstwa domowe o niskich i średnich dochodach. Zielone linie oznaczają przedziały ufności dla zebranych danych empirycznych na poziomie 95%.



Rysunek 3.11: Dopasowanie uogólnionego modelu Lotka–Volterra (linia ciągła) do empirycznego dopełnienia dystrybucyjności rocznych dochodów rozporządzalnych gospodarstw domowych w Polsce (punkty) z lat 2007–2009. Ponownie widzieć, że model ten dobrze opisuje gospodarstwa domowe o niskich i średnich dochodach. Zielone linie oznaczają przedziały ufności dla zebranych danych empirycznych na poziomie 95%.



Rysunek 3.12: Dopasowanie uogólnionego modelu Lotka–Volterra (linia ciągła) do empirycznego dopełnienia dystrybuanty rocznych dochodów rozporządzalnych gospodarstw domowych w Polsce (punkty) z roku 2010. Nadal widać, że model ten dobrze opisuje gospodarstwa domowe o niskich i średnich dochodach. Zielone linie oznaczają przedziały ufności dla zebranych danych empirycznych na poziomie 95%.

Uogólniony model Lotka–Volterra dobrze opisuje gospodarstwa domowe uzyskujące niskie i średnie dochody. Istotną zaletą tego modelu jest możliwość scharakteryzowania dopełnień dystrybuant rocznych dochodów rozporządzalnych gospodarstw domowych przy pomocy pojedynczej formy funkcyjnej. Na wykresach 3.10 – 3.12 zaznaczono również przedziały ufności (zielone linie) dla zebranych danych empirycznych na poziomie 95% (na podstawie pracy [99]). Ponownie widać, że niektóre z punktów empirycznego dopełnienia dystrybuanty leżą poza wyznaczonymi przedziałami ufności. Zatem mają miejsce odstępstwa od prawa potęgowego. Wartości parametrów dla poszczególnych dopasowań zamieszczono w tabeli 3.3¹¹.

¹¹ Błędy parametrów zostały zamieszczone w legendach na rysunkach 3.10 – 3.12. Z kolei, wartość średnią $\langle x \rangle$ oraz rozpiętość R_x można oszacować na podstawie uzyskanych wartości parametru α .

Tabela 3.3: Parametry α uzyskane z dopasowania uogólnionego modelu Lotka–Volterra do empirycznego dopełnienia dystrybucyjności rocznych dochodów rozporządzalnych gospodarstw domowych w Polsce z lat 2004–2010. W przedostatnim wierszu znajduje się wartość średnia parametru α . Możemy zauważyć, że jego rozrzut względny (względna rozpiętość) R_x wynosi 6%.

Rok	α
2004	3.776
2005	3.766
2006	3.720
2007	3.572
2008	3.546
2009	3.602
2010	3.558
$\langle x \rangle$	3.649
$R_x = \frac{x_{\max} - x_{\min}}{\langle x \rangle}$	0.06

W przypadku uogólnionego modelu Lotka–Volterra można wyznaczyć efektywną minimalną roczną wartość dochodu rozporządzalnego gospodarstwa domowego. Jest to wartość, poniżej której prawdopodobieństwo osiągnięcia przez gospodarstwa domowe mniejszych dochodów gwałtownie maleje (rozkład gęstości prawdopodobieństwa uogólnionego modelu Lotka–Volterra gwałtownie maleje do zera):

$$\frac{m_{\min}}{\langle m \rangle} = \frac{\alpha - 1}{\alpha + 1}. \quad (3.4)$$

Na podstawie powyższego wzoru otrzymujemy minimalne wartości rocznego dochodu – tabela 3.4¹².

¹² Wartości wyznaczone z dokładnością do 10 groszy.

Tabela 3.4: Średnie oraz minimalne roczne wartości dochodu rozporządzalnego gospodarstw domowych w Polsce w latach 2004–2010.

Rok	$\langle m \rangle$ [zł]	m_{\min} [zł]
2004	26 744.0	15 544.7
2005	27 350.3	15 873.0
2006	29 540.6	17 023.4
2007	32 674.4	18 381.1
2008	36 556.7	20 473.7
2009	38 713.9	21 889.1
2010	41 275.8	23 164.4

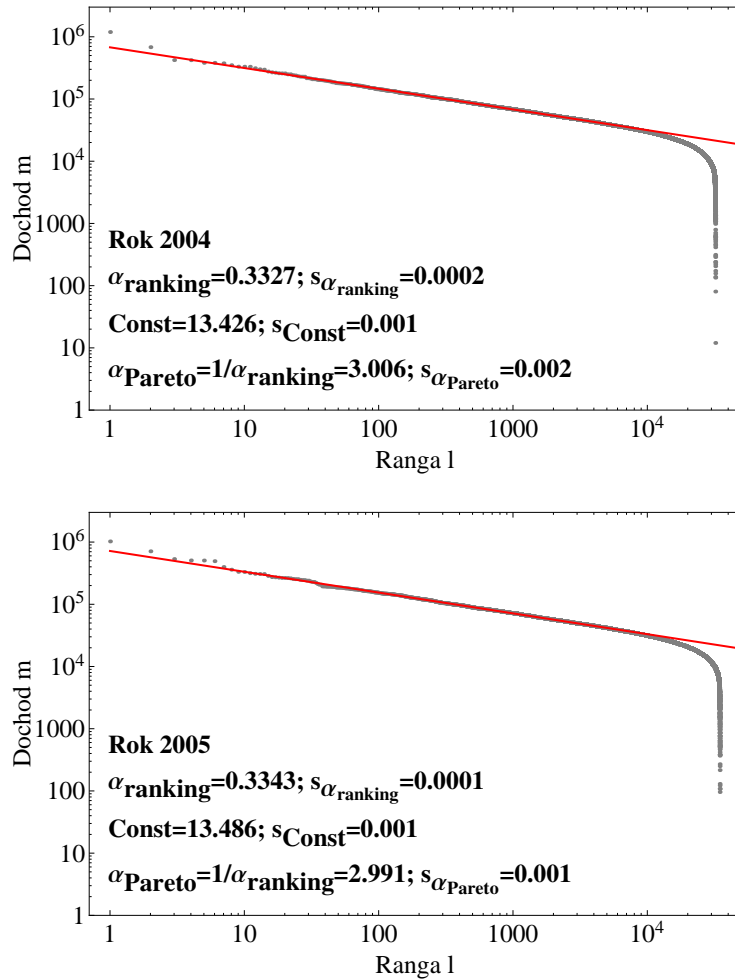
Niestety, baza danych Głównego Urzędu Statystycznego jaką dysponuję nie pozwala na rozstrzygnięcie pomiędzy Prawem Efektów Proporcjonalnych i słabym prawem Pareto z jednej strony a uogólnionym modelem Lotka–Volterra z drugiej. Zatem, na przyjętym poziomie dokładności mogą być stosowane oba podejścia.

Rankingi gospodarstw domowych o niskich i średnich dochodach

W dalszej części rozważań zostanie przedstawiona alternatywna metoda wyznaczania wykładnika Pareto na podstawie danych empirycznych dotyczących gospodarstw domowych zarówno o średnich jak i wysokich dochodach.

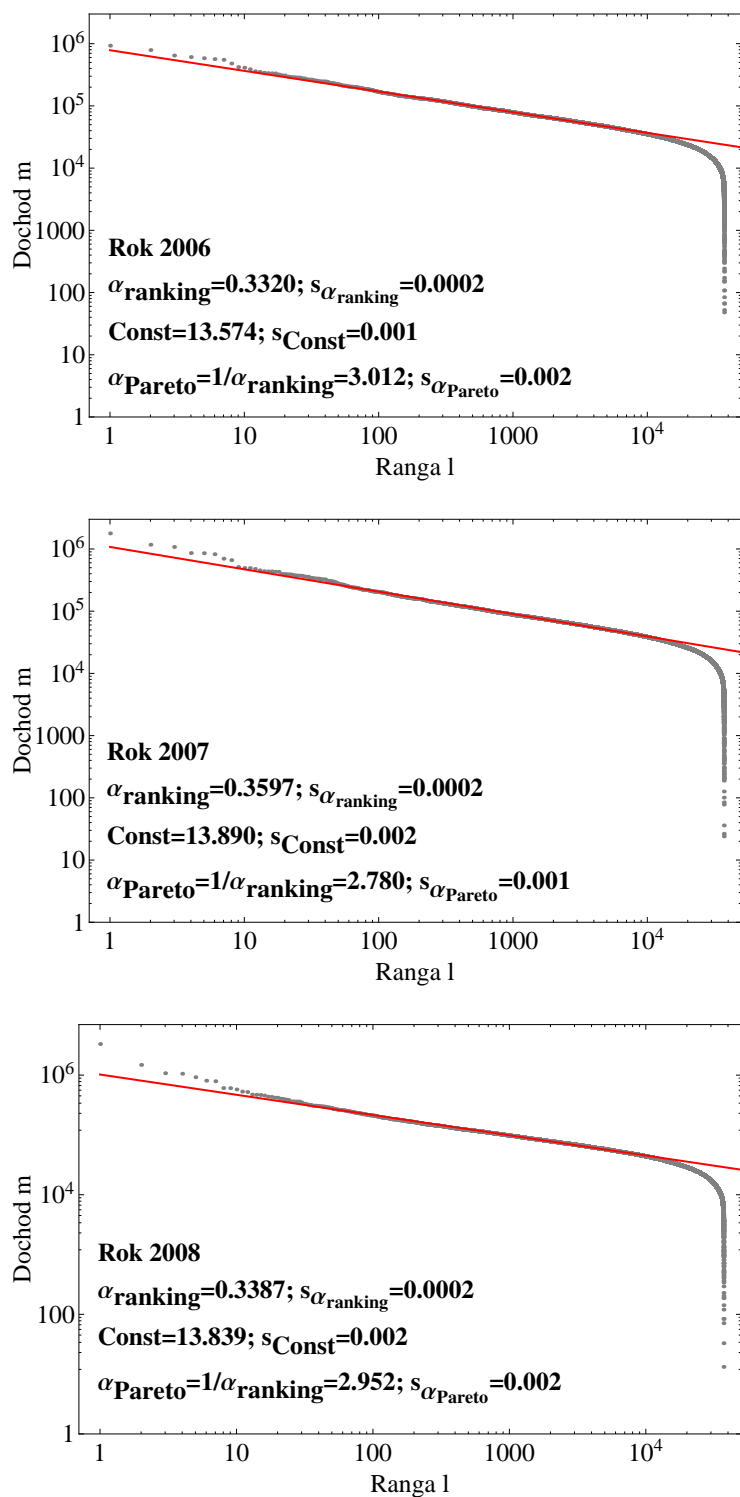
Jak to zostało wspomniane wcześniej, wyznaczenie wykładnika Pareto na podstawie empirycznych dopełnień dystrybuant dla gospodarstw domowych osiągających wysokie dochody nie jest możliwe, gdyż rozkłady te zawierają zbyt mało danych empirycznych (przypomnijmy, że zostały one zbudowane na bazie danych empirycznych Głównego Urzędu Statystycznego). Wartość wykładnika Pareto otrzymamy analizując ranking gospodarstw domowych, czyli wykres dochodów gospodarstw domowych w zależności od ich miejsca w rankingu (dochody zostały uporządkowane w sposób malejący – zatem pierwsze miejsce w rankingu zajmuje najbogatsze w danej próbie gospodarstwo domowe) [90]. Jest to podejście równoważne analizie empirycznych dopełnień dystrybuant. Różnica polega na in-

nym nazewnictwie i innej reprezentacji skal. Zwyczajowo w badaniach analizuje się albo dopełnienia dystrybuant, albo rankingi. Odpowiednie wykresy w skali logarytmicznej zamieszczono się na rysunkach 3.13 – 3.15¹³.

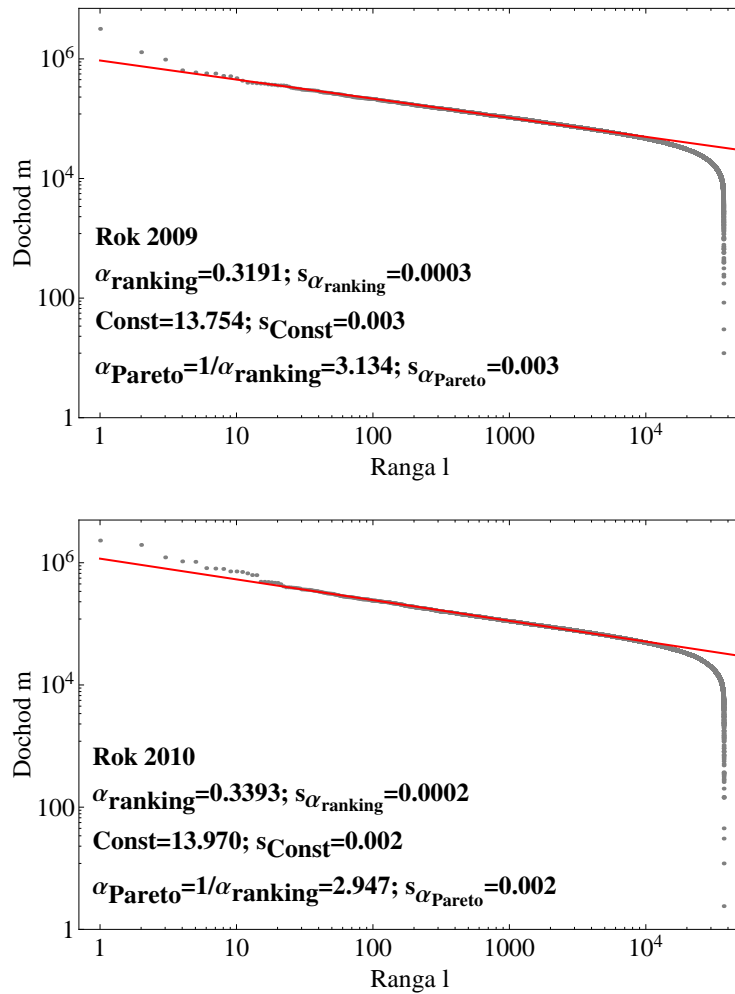


Rysunek 3.13: Ranking gospodarstw domowych o niskich i średnich dochodach (linia ciągła oznacza dopasowanie, a punkty dane empiryczne) – lata 2004–2005. Jak widać, gospodarstwa domowe o średnich dochodach są całkiem dobrze opisywane prostą o nachyleniu $-\alpha_{\text{ranking}}$.

¹³ Na rysunkach 3.13 – 3.18 wielkość Const pochodzi z dopasowania prostej (w skali log-log). Dopasowując do danych empirycznych (dotyczących średnich dochodów gospodarstw domowych w Polsce) prostą w skali log-log wyboru podpóbki dokonano analogicznie jak w przypadku słabego prawa Pareto.



Rysunek 3.14: Ranking gospodarstw domowych o niskich i średnich dochodach (linia ciągła oznacza dopasowanie, a punkty dane empiryczne) – lata 2006–2008. Jak widać, gospodarstwa domowe o średnich dochodach są całkiem dobrze opisywane prostą o nachyleniu $-\alpha_{\text{ranking}}$.



Rysunek 3.15: Ranking gospodarstw domowych o niskich i średnich dochodach (linia ciągła oznacza dopasowanie, a punkty dane empiryczne) – lata 2009–2010. Jak widać, gospodarstwa domowe o średnich dochodach są całkiem dobrze opisywane prostą o nachyleniu $-\alpha_{\text{ranking}}$.

Gospodarstwa domowe o średnich dochodach opisywane są w skali logarytmicznej prostą o nachyleniu $-\alpha_{\text{ranking}}$. Jak należało oczekiwać, odwrotność α_{ranking} jest tutaj zbliżona do wyznaczonych wartości wykładnika Pareto, α_{Pareto} , uzyskanych w przypadku dopełnień dystrybuant rocznych dochodów rozporządzalnych gospodarstw domowych. Wskazuje to na zgodność obu zastosowanych metod [90]. Porównanie wartości wykładników Pareto uzyskanych na obu drogach zamieszczono w tabeli 3.5¹⁴.

¹⁴ Błędy wykładników zostały zamieszczone w legendach na rysunkach 3.7 – 3.9 oraz 3.13 – 3.15. Z kolei, wartość średnią $\langle x \rangle$ oraz rozpiętość R_x można oszacować na podstawie uzyskanych wartości wykładników α_{Pareto} .

Tabela 3.5: Porównanie wartości wykładników Pareto uzyskanych z analizy rankingów dochodów gospodarstw domowych z wartościami wykładników Pareto uzyskanych z dopasowania słabego prawa Pareto do empirycznego dopełnienia dystrybuanty rocznych dochodów rozporządzalnych gospodarstw domowych w Polsce w latach 2004–2010. W przedostatnim wierszu znajdują się wartości średnie wykładników Pareto. Możemy zauważyć, że rozrzut względny (względna rozpiętość) R_x wykładników nie przekracza 12%.

Rok	$\alpha_{\text{Pareto}} \text{ (ranking)}$	$\alpha_{\text{Pareto}} \text{ (dopełnienie dystrybuanty)}$
2004	3.006	3.005
2005	2.991	3.009
2006	3.012	2.984
2007	2.780	2.809
2008	2.952	3.146
2009	3.134	3.124
2010	2.947	3.095
$\langle x \rangle$	2.975	3.025
$R_x = \frac{x_{\max} - x_{\min}}{\langle x \rangle}$	0.12	0.11

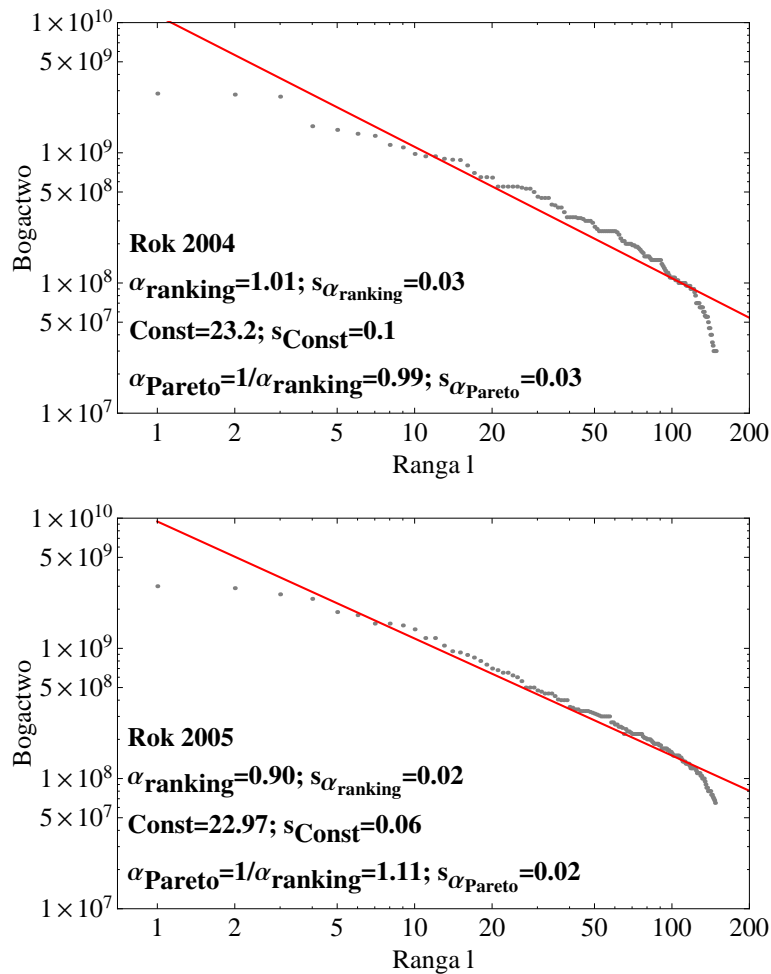
Rankingi gospodarstw domowych o wysokich dochodach

W dotychczasowych rozważaniach nie analizowano dochodów gospodarstw domowych o wysokich dochodach, gdyż obserwacje takie były zbyt mało liczne, aby można było je poddać jakimkolwiek opisowi statystycznemu. Problem ten można pośrednio rozwiązać analizując wartość majątku najbogatszych Polaków (ok. 100–150 obserwacji) w postaci rankingu [101] (rysunki 3.16 – 3.18, skala logarytmiczna). Zgodnie z tym, co wykazano wcześniej, jeśli ranking najbogatszych można w skali logarytmicznej opisać zależnością liniową¹⁵, to wówczas podlegają oni słabemu prawu Pareto o wykładniku¹⁶ $\alpha_{\text{Pareto}} = \frac{1}{\alpha_{\text{ranking}}}$. Choć w tym przypadku zamiast o dochodach mówię o bogactwie, to jednak zakładam tutaj, że

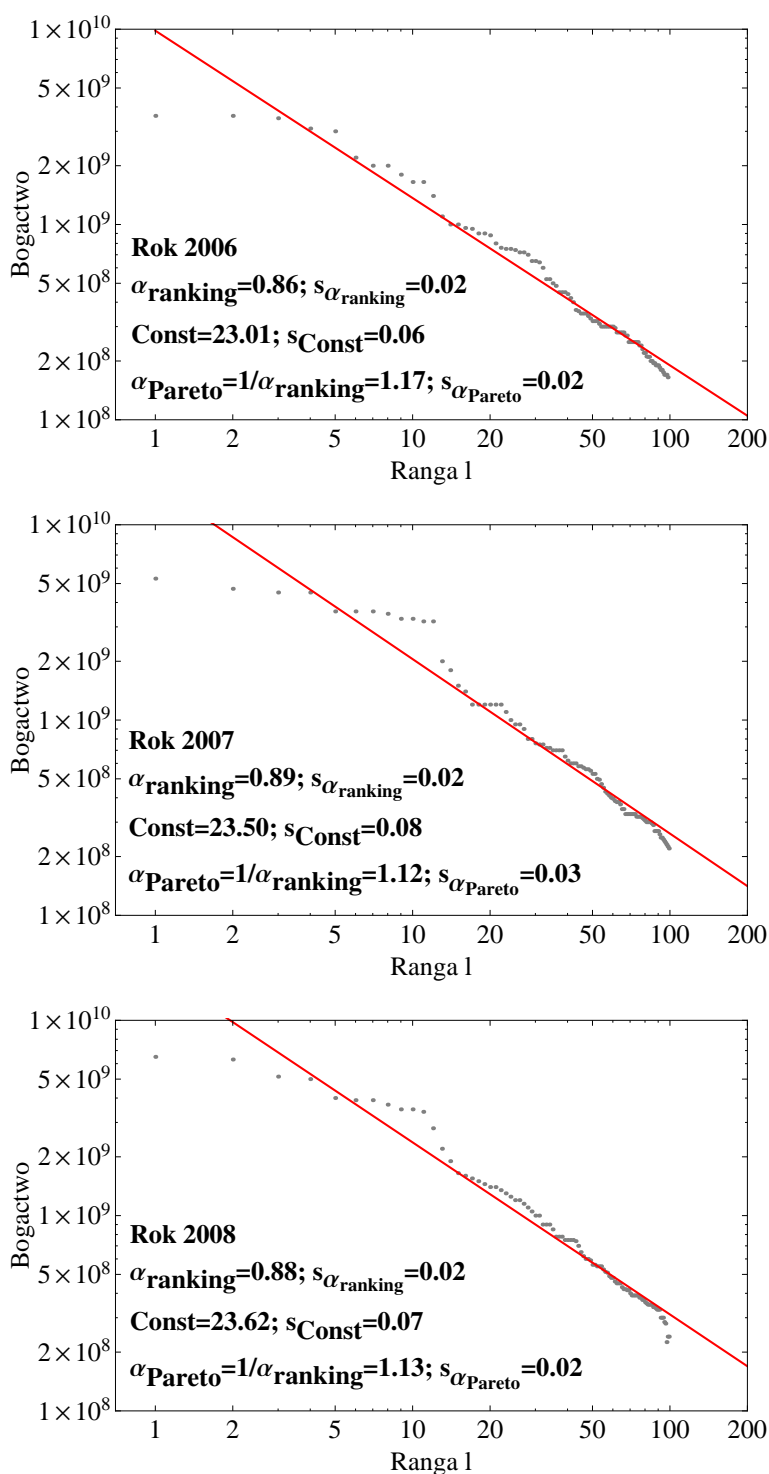
¹⁵ Dopasanie prostej w skali log-log wykonano na całej próbie.

¹⁶ Pojawiającego się tutaj wykładnika α_{Pareto} odnoszącego się do majątku najbogatszych Polaków, nie należy mylić z wykładnikiem α_{Pareto} odnoszącym się do gospodarstw domowych o średnich dochodach.

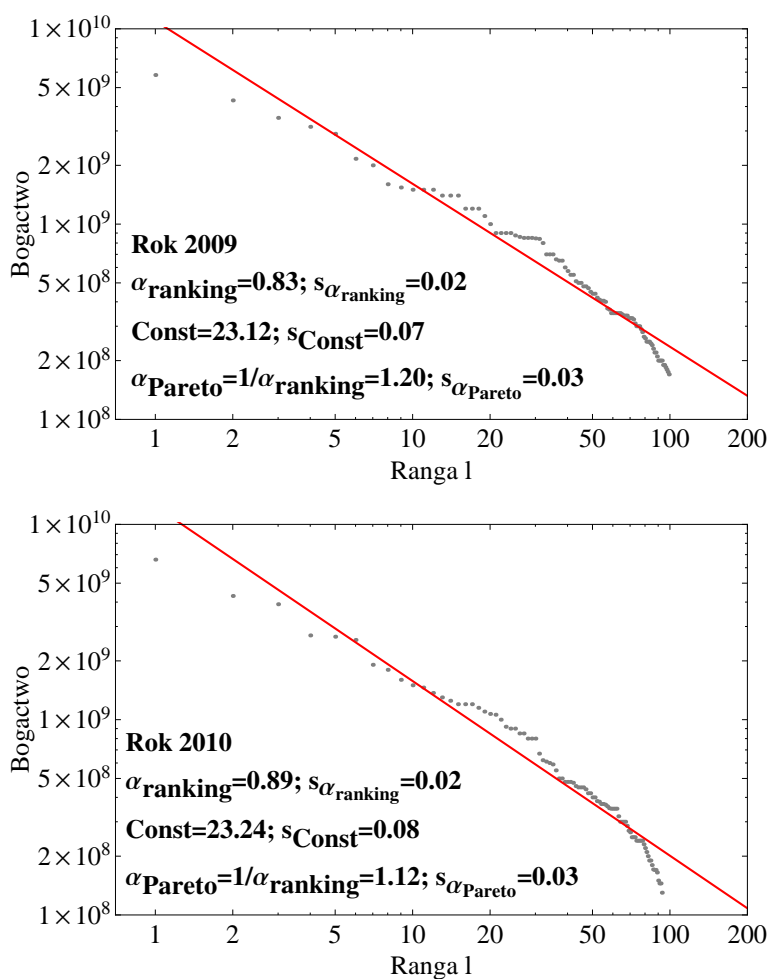
z roku na rok nie zmienia się klasa uniwersalności (wykładnik Pareto zmienia się nieznacznie – jest wolnozmienną funkcją lat), a stąd wykładnik ten może również opisywać wysokie dochody gospodarstw domowych [4, 12].



Rysunek 3.16: Ranking najbogatszych Polaków (linia ciągła oznacza dopasowanie, a punkty dane empiryczne) – lata 2004–2005. Jak widać, gospodarstwa domowe o wysokich dochodach są całkiem dobrze opisywane prostą o nachyleniu $-\alpha_{\text{ranking}}$.



Rysunek 3.17: Ranking najbogatszych Polaków (linia ciągła oznacza dopasowanie, a punkty dane empiryczne) – lata 2006–2008. Jak widać, gospodarstwa domowe o wysokich dochodach są całkiem dobrze opisywane prostą o nachyleniu $-\alpha_{\text{ranking}}$.



Rysunek 3.18: Ranking najbogatszych Polaków (linia ciągła oznacza dopasowanie, a punkty dane empiryczne) – lata 2009–2010. Jak widać, gospodarstwa domowe o wysokich dochodach są całkiem dobrze opisywane prostą o nachyleniu $-\alpha_{\text{ranking}}$.

Wartości wykładników Pareto pochodzące z poszczególnych dopasowań zostały zebrane w tabeli 3.6¹⁷.

¹⁷ Błędy wykładników zostały zamieszczone w legendach na rysunkach 3.16 – 3.18. Z kolei, średnią $\langle x \rangle$ oraz rozpiętość R_x można oszacować na podstawie uzyskanych wartości wykładników α_{Pareto} .

Tabela 3.6: Wykładniki Pareto uzyskane z analizy rankingów najbogatszych Polaków z lat 2004–2010. W przedostatnim wierszu znajduje się wartość średnia wykładnika Pareto. Możemy zauważyć, że jego rozrzut względny (względna rozpiętość) R_x wynosi 19%.

Rok	α_{Pareto}
2004	0.99
2005	1.11
2006	1.17
2007	1.12
2008	1.13
2009	1.20
2010	1.12
$\langle x \rangle$	1.12
$R_x = \frac{x_{\max} - x_{\min}}{\langle x \rangle}$	0.19

Otrzymane wartości wykładnika Pareto są bliskie jedności, co jest zgodne z podobnymi analizami dotyczącymi majątku najbogatszych w innych krajach [2,3,5,30,44,67,102,103]. Zatem można oczekiwać, że mechanizm zdobywania wysokich dochodów przez gospodarstwa domowe, opiera się na istnieniu konkurencji rynkowej (podobnie jak w przypadku gospodarstw domowych o średnich dochodach) [100], jednak w tym wypadku jednostki tworzące gospodarstwa domowe są przeważnie właścicielami konkurujących firm, przy czym dochody tych firm opisywane są prawem Zipfa¹⁸ [88,91,92].

Ponieważ dla gospodarstw domowych o wysokich dochodach dysponujemy niewielką liczbą obserwacji (ok. 100–150), mamy do czynienia ze znaczącym rozrzutem danych, co wpływa na dokładność obliczonych wykładników Pareto. Dokładność ta jest o rząd wielkości mniejsza, niż w przypadku gospodarstw domowych o średnich dochodach.

¹⁸ Prawo Zipfa jest prawem Pareto z wykładnikiem $\alpha = 1$.

3.4 PODSUMOWANIE ROZDZIAŁU

W niniejszym rozdziale dokonano analizy empirycznych dopełnień dystrybucji dotyczących rocznych dochodów rozporządzalnych gospodarstw domowych w Polsce w latach 2004–2010. Okazało się, że mogą one być dobrze opisywane dopełnieniem dystrybucji rozkładu log-normalnego w przypadku gospodarstw domowych o niskich dochodach, jak również słabym prawem Pareto dla gospodarstw uzyskujących średnie dochody. Wśród rozważanych modeli obiecujące wyniki uzyskano także dla uogólnionego modelu Lotka–Volterra. Model ten w dobrym przybliżeniu odzwierciedla analizowane rozkłady oraz oferuje cenne podejście teoretyczne na poziomie mikroskopowym, w którym dochód gospodarstwa domowego determinowany jest przez dochody dotychczas uzyskiwane oraz przez wysokość świadczeń pomocy społecznej (ogólniej mówiąc redystrybucję dochodów w społeczeństwie) i ogólny stan gospodarki. Dzięki temu, iż wykładnik Pareto opisujący gospodarstwa domowe o średnich dochodach, można uzyskać zarówno poprzez analizę dopełnień dystrybucji jak i rankingu, udało się wyznaczyć także jego wartość dla najbogatszych jednostek społeczeństwa polskiego.

Wyznaczone parametry pochodzące z różnych modeli mogą, jak przypuszczam, stanowić wskaźniki, czy nawet prekursorzy kryzysu. Kryzys nie wpływa na gospodarstwa domowe o niskich dochodach (parametry dopełnienia dystrybucji rozkładu log-normalnego praktycznie nie zmieniają się), natomiast prowadzi do mniejszego rozwarstwienia społecznego wśród gospodarstw domowych o średnich i wysokich dochodach (oba wykładniki Pareto rosną). Warto jednak zauważyć, że w przypadku gospodarstw domowych o wysokich dochodach wahania wykładnika Pareto są stosunkowo niewielkie co oznacza, że skutki finansowe kryzysu dla najbogatszych obywateli społeczeństwa polskiego nie były aż tak dotkliwe, jak w innych krajach. Jednak w celu wyprowadzenia dalszych wniosków niezbędne są tutaj dodatkowe badania, polegające przede wszystkim na przeprowadzeniu porównań z wcześniejszymi kryzysami.

Klasa społeczna reprezentowana przez gospodarstwa domowe o niskich dochodach wydaje się niezwykle stabilna. Jej dochody nie ulegają znaczącym zmianom w wyniku kryzysu. Zatem parametry rozkładów opisujących niskie dochody go-

spodarstw domowych nie mogą być indykatorami kryzysu. Tylko analiza średnich i wysokich dochodów gospodarstw domowych może dać odpowiedź na pytanie czy nadchodzi kryzys.

Zatem, roczne dochody rozporządzalne polskich gospodarstw domowych mogą być opisane za pomocą rozkładów:

- dopełnienia dystrybuanty rozkładu log-normalnego, w przypadku gospodarstw domowych o niskich dochodach,
- dopełnienia dystrybuanty rozkładu Pareto z wykładnikiem około 3, w przypadku gospodarstw domowych o średnich dochodach,
- dopełnienia dystrybuanty rozkładu Pareto z wykładnikiem około 1, w przypadku gospodarstw domowych o wysokich dochodach.

ANALIZA ROCZNYCH DOCHODÓW GOSPODARSTW DOMOWYCH W UNII EUROPEJSKIEJ

„Nauka jest zbiorem wypróbowanych przepisów.”

PAUL AMBROISE VALÉRY

4.1 OPIS BAZ DANYCH

W analizie danych empirycznych wykorzystano dane Eurostatu pochodzące z badania EU-SILC [104–111] (Survey on Income and Living Conditions) z lat 2004–2010, czyli z Europejskiego Badania Warunków Życia Ludności, obejmującego ogólne informacje o cechach demograficznych gospodarstw domowych w Unii Europejskiej, ich warunkach życia, poziomie dochodów oraz aktywności ekonomicznej. W badaniach użyto także danych o zamożności najbogatszych Europejczyków w latach 2003–2010 pochodzące z rankingu magazynu Forbes – The World’s Billionaires [112].

W przypadku danych EU-SILC dochód, który został poddany analizie to roczny dochód brutto gospodarstwa domowego¹ (ang. *Total household gross income*). Zgodnie z definicją (tłumaczenie za [104]), roczny dochód brutto gospodarstwa domowego to „...pieniężny i niepieniężny dochód uzyskany przez gospodarstwo domowe w okresie jednego roku przed odliczeniem podatków dochodowych, podat-

¹ Należy zaznaczyć, że w bazie Eurostatu znajduje się również zmienna – roczny dochód brutto do dyspozycji gospodarstwa domowego (ang. *Total gross disposable household income*). Nie została ona jednak wybrana do analizy, gdyż sposób jej obliczania może być różny w zależności od kraju (głównie ze względu na różne systemy podatkowe i ubezpieczeń społecznych).

ków od dochodów z własności, składek na ubezpieczenie społeczne i zdrowotne oraz uwzględniający przepływy pieniężne otrzymane od innych gospodarstw domowych”².

Z bazy danych EU-SILC usunięto rekordy, dla których dochód był mniejszy bądź równy zero. Niedodatnie wartości dochodu uznano za obserwacje błędne. Podsumowanie (w rozbiciu na poszczególne lata) liczby obserwacji w analizowanej próbie, jak i krajów, dla których dostępne były dane o dochodach, zostało zamieszczone w tabeli 4.1³.

² Dokładny opis konstrukcji zmiennej – roczny dochód brutto gospodarstwa domowego (*Total household gross income*) został przedstawiony w dodatku E.

³ W analizie danych empirycznych uwzględniono dane o dochodach pochodzące z krajów należących do Unii Europejskiej (UE), jak również dane o dochodach pochodzące z krajów nie należących do UE, które jednak biorą udział w badaniu EU-SILC.

Tabela 4.1: Kraje należące do Unii Europejskiej oraz kraje spoza Unii Europejskiej, dla których w bazie EU-SILC, dla poszczególnych lat dostępne są dane o rocznych dochodach brutto gospodarstw domowych. W tabeli podano również liczbę obserwacji wykorzystanych w analizie danych empirycznych.

Rok	Liczba obserwacji	Kraje (UE)	Kraje (spoza UE)
2004	55 477	Austria, Belgia, Dania, Estonia, Finlandia, Irlandia, Luksemburg, Szwecja	Islandia, Norwegia
2005	138 593	Austria, Belgia, Cypr, Czechy, Dania, Estonia, Finlandia, Holandia, Irlandia, Litwa, Luksemburg, Niemcy, Polska, Portugalia, Słowacja, Słowenia, Szwecja, Węgry, Wielka Brytania	Islandia, Norwegia
2006	156 734	Austria, Belgia, Cypr, Czechy, Dania, Estonia, Finlandia, Hiszpania, Holandia, Irlandia, Litwa, Luksemburg, Niemcy, Polska, Słowacja, Słowenia, Szwecja, Węgry, Wielka Brytania	Islandia, Norwegia
2007	206 488	Austria, Belgia, Cypr, Czechy, Dania, Estonia, Francja, Finlandia, Grecja, Hiszpania, Holandia, Irlandia, Litwa, Luksemburg, Łotwa, Niemcy, Polska, Portugalia, Słowacja, Słowenia, Szwecja, Węgry, Wielka Brytania, Włochy	Islandia, Norwegia
2008	209 017	Austria, Belgia, Bułgaria, Cypr, Czechy, Dania, Estonia, Finlandia, Grecja, Hiszpania, Holandia, Irlandia, Litwa, Luksemburg, Łotwa, Niemcy, Polska, Portugalia, Rumunia, Słowacja, Słowenia, Szwecja, Węgry, Wielka Brytania, Włochy	Islandia, Norwegia
2009	223 882	Austria, Belgia, Bułgaria, Cypr, Czechy, Dania, Estonia, Finlandia, Francja, Grecja, Hiszpania, Holandia, Irlandia, Litwa, Luksemburg, Łotwa, Malta, Niemcy, Polska, Portugalia, Rumunia, Słowacja, Słowenia, Szwecja, Węgry, Wielka Brytania, Włochy	Islandia, Norwegia
2010	216 873	Austria, Belgia, Bułgaria, Czechy, Dania, Estonia, Finlandia, Francja, Grecja, Hiszpania, Holandia, Irlandia, Litwa, Luksemburg, Łotwa, Malta, Niemcy, Polska, Portugalia, Rumunia, Słowacja, Słowenia, Szwecja, Węgry, Wielka Brytania, Włochy	Islandia, Norwegia

Dane Eurostatu zawierają dla poszczególnych lat tylko kilka obserwacji dotyczących gospodarstw domowych o wysokich dochodach, co powoduje, że nie można poddać ich jakimukolwiek opisowi statystycznemu. Zatem, aby dokonać analizy gospodarstw domowych o wysokich dochodach wykorzystano dane magazynu Forbes dotyczące rankingu majątków posiadanych przez najbogatszych ludzi na świecie – „The World’s Billionaires list” [112]. W zestawieniu tym zamieszczane są osoby, których wartość majątku w danym roku przekroczyła miliard dolarów. Z listy tej wybrano tych, którzy przebywali w danym roku na terenie Unii Europejskiej i na tej podstawie utworzono rankingi najbogatszych Europejczyków w latach 2004–2010. Należy zaznaczyć, że nawet jeśli jakiś kraj nie dostarczył danych do EU-SILC dla wybranego roku to i tak został on uwzględniony przy tworzeniu rankingu. W analizie wzięto również pod uwagę osoby przebywające na terenie Islandii i Norwegii, ponieważ kraje te udostępniają dane o dochodach gospodarstw domowych na potrzeby badania EU-SILC.

Baza danych magazynu Forbes będzie przekształcana i łączona z bazą danych EU-SILC dlatego też procedura tego łączenia zostanie przedstawiona w rozdziale 4.2.

4.2 PROCEDURA ŁĄCZENIA BAZ DANYCH

Jak wspomniano wcześniej dane EU-SILC zawierają niewielką liczbę obserwacji dotyczących gospodarstw domowych o wysokich dochodach. Aby możliwa była systematyczna analiza dochodów wszystkich trzech klas społecznych należało utworzyć poszerzoną bazę danych zawierającą informacje o gospodarstwach domowych uzyskujących niskie, średnie i wysokie dochody. W tym celu połączono bazę EU-SILC i bazę magazynu Forbes. Procedura łączenia baz danych przebiegała następująco:

1. Najpierw oszacowano dochody uzyskiwane przez najbogatszych Europejczyków. Aby dla konkretnej osoby z rankingu Forbesa uzyskać jej dochód w danym roku, należało obliczyć różnicę między jej majątkiem w tym roku i w roku poprzednim. Inaczej mówiąc, przyjęto założenie, że dochód najbogatszych

jest proporcjonalny do różnicy ich majątku między dwoma kolejnymi latami. Oczywiście w dalszej analizie uwzględniono tylko te osoby, dla których różnica majątków była dodatnia, zaniebując tych którzy odnieśli straty, bądź wartość ich majątku nie uległa zmianie.

2. Wartość majątku podawana jest przez redaktorów magazynu Forbes w dolarach amerykańskich. Zatem, po obliczeniu wartości uzyskiwanego dochodu konieczne było jego przeliczenie na euro. Kurs wymiany⁴ dolara na euro dla poszczególnych lat został wzięty z tego samego dnia, dla którego dokonywali przeliczeń majątków redaktorzy Forbesa (patrz tabela 4.2).
3. W kolejnym kroku, mając tak przekształcone dane Forbesa, połączono je wprost z danymi EU-SILC. Na podstawie uzyskanej bazy danych o zwiększonej w ten sposób liczebności, zawierającej informacje o dochodach wszystkich trzech klas społecznych, skonstruowano empiryczne dopełnienia dystrybuant zgodnie z regułami opisanymi w rozdziale 3.2. Niestety, rozkłady te charakteryzowały się dużą przerwą (brakiem danych) pomiędzy gospodarstwami domowymi uzyskującymi wysokie dochody – pochodzącymi z bazy danych Eurostatu, a gospodarstwami domowymi uzyskującymi wysokie dochody – pochodzącymi z bazy danych magazynu Forbes.
4. Zakładając, że statystyki budowane na bazie dochodu (patrz rozdział 3.2) są ciągłą funkcją tegoż dochodu, usunięto wspomnianą przerwę. W tym celu dochody gospodarstw domowych pochodzących z rankingu Forbesa pomnożono przez jednakowy czynnik proporcjonalności, który wyniósł 1.0×10^{-2} dla wszystkich analizowanych lat. Pomnożenie przez ten stały czynnik odpowiada właściwemu doborowi stałej proporcjonalności; konieczność znalezienia takiej stałej wynika z rozważań znajdujących się w punkcie 1 (gdyż różnica majątków nie stanowi bezpośrednio dochodu, ale pomiędzy dochodem a różnicą majątków istnieje proporcjonalność). Uzyskano zatem, empiryczne dopełnienia dystrybuant dla lat 2004–2010 pozbawione jakichkolwiek przerw. Zastosowana modyfikacja doprowadziła do niewielkiego przekrywania się

⁴ Kursy wymiany walut zostały zaczerpnięte ze strony Europejskiego Banku Centralnego – <http://www.ecb.int/>

(w obszarze wspólnym) danych empirycznych z obu baz (patrz rozdział 4.3, rysunki 4.1 – 4.3), co potwierdza poprawność zastosowanej procedury.

Zaproponowana przeze mnie procedura łączenia baz danych jest jednym z kluczowych elementów niniejszej rozprawy doktorskiej. Doprowadziła ona do uzyskania wystarczającej liczby danych empirycznych dla wszystkich trzech klas społecznych, tj. gospodarstw domowych uzyskujących niskie, średnie i wysokie dochody. Było to możliwe dzięki odpowiedniemu zmodyfikowaniu danych dotyczących wartości majątków najbogatszych Europejczyków i połączeniu ich z danymi o dochodach pochodzącymi z bazy EU-SILC. Chociaż dane magazynu Forbes stanowią przybliżone oszacowanie majątków posiadanych przez najbogatszych, to pozwalają one dość dobrze ustalić ich hierarchię wyrażoną poprzez ranking dochodów. Fakt ten w wystarczającym stopniu uzasadnia zaproponowaną metodę, gdyż głównym jej celem jest zaliczenie gospodarstw domowych o wysokich dochodach do konkretnej klasy uniwersalności (indeksowanej danym wykładnikiem Pareto), a nie dokładne wyznaczenie wielkości ich całkowitego dochodu.

Liczba obserwacji w analizowanej próbie, kurs wymiany wraz z datą oraz kraje uwzględnione w połączonej bazie danych zostały zamieszczone w tabeli 4.2.

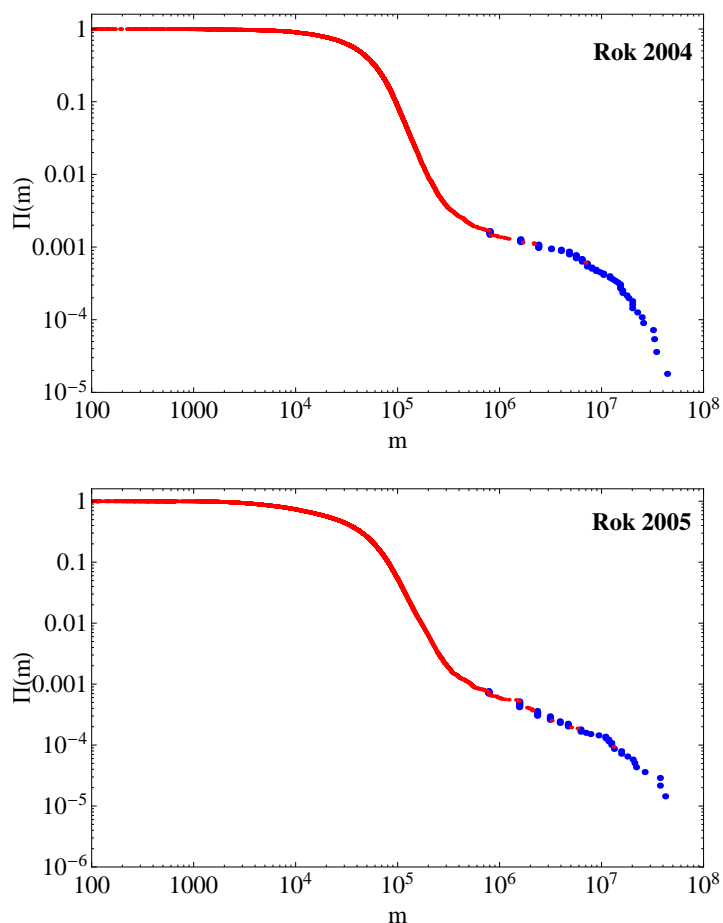
Tabela 4.2: Zbiorcze podsumowanie bazy danych powstałej z połączenia danych EU-SILC i magazynu Forbes.

Rok	Liczba obserwacji	Kurs wymiany, data	Kraje
2004	55 554 (w tym 77 z Forbesa)	1 USD = 0.7981 EUR, 6 luty 2004	EU-SILC 2004, plus: Francja, Hiszpania, Holandia, Niemcy, Portugalia, Wielka Brytania, Włochy
2005	138 664 (w tym 71 z Forbesa)	1 USD = 0.7779 EUR, 11 luty 2005	EU-SILC 2005, plus: Francja, Hiszpania, Włochy
2006	156 810 (w tym 76 z Forbesa)	1 USD = 0.8412 EUR, 13 luty 2006	EU-SILC 2006, plus: Francja, Portugalia, Włochy
2007	206 608 (w tym 120 z Forbesa)	1 USD = 0.7688 EUR, 9 luty 2007	EU-SILC 2007
2008	209 115 (w tym 98 z Forbesa)	1 USD = 0.6877 EUR, 11 luty 2008	EU-SILC 2008, plus: Francja
2009	223 885 (w tym 3 z Forbesa)	1 USD = 0.7798 EUR, 13 luty 2009	EU-SILC 2009
2010	216 983 (w tym 110 z Forbesa)	1 USD = 0.7368 EUR, 12 luty 2010	EU-SILC 2010

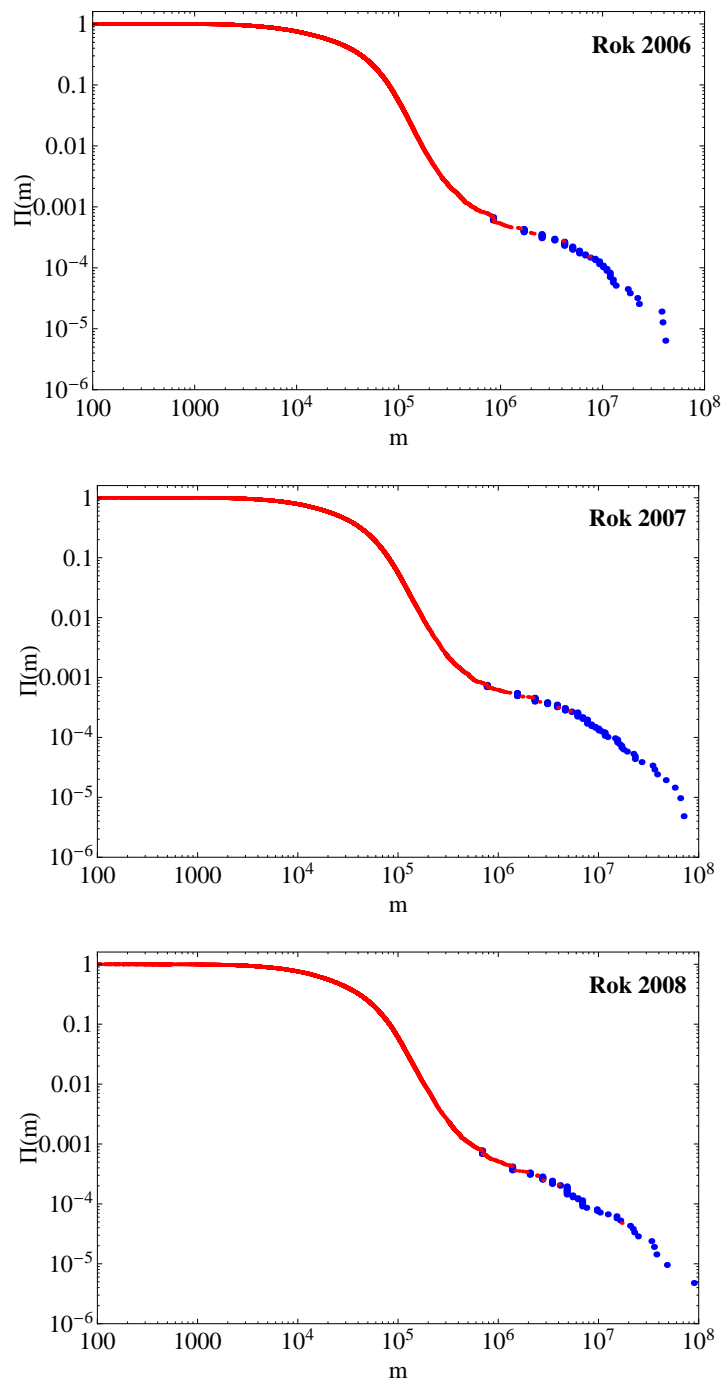
4.3 EMPIRYCZNE DOPEŁNIENIA DYSTRYBUANT

Jednym z celów niniejszej rozprawy doktorskiej jest analiza empirycznych dopełnień dystrybuant rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej. Rozkłady te skonstruowano zgodnie z regułami opisanymi w rozdziale 3.2 przy użyciu bazy danych, której proces tworzenia został omówiony w rozdziale 4.2.

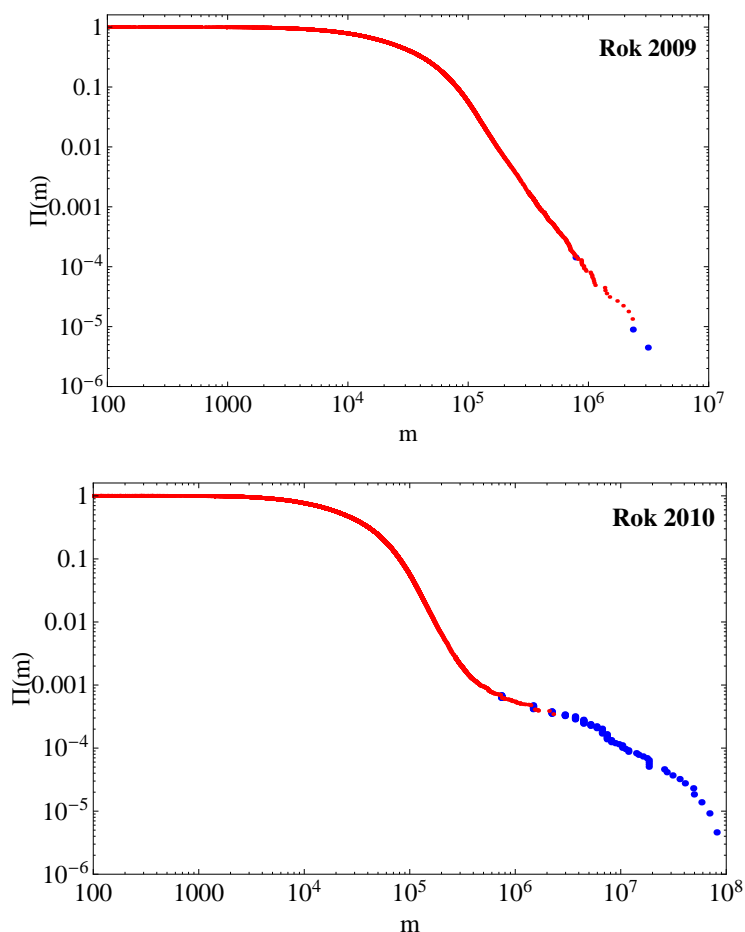
Empiryczne dopełnienia dystrybuant rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej zostały przedstawione (w skali logarytmicznej) na rysunkach 4.1 – 4.3.



Rysunek 4.1: Empiryczne dopełnienia dystrybuant, $\Pi(m)$, rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej, m , z lat 2004–2005. Punkty czerwone reprezentują dane empiryczne pochodzące z badania EU-SILC a niebieskie dane empiryczne pochodzące z magazynu Forbes. Widać, że dane z obszaru wspólnego dochodów przekrywają się, jak być powinno.



Rysunek 4.2: Empiryczne dopełnienia dystrybuant, $\Pi(m)$, rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej, m , z lat 2006–2008. Punkty czerwone reprezentują dane empiryczne pochodzące z badania EU-SILC a niebieskie dane empiryczne pochodzące z magazynu Forbes. Widać (podobnie jak dla lat 2004–2005), że dane ze wspólnego obszaru dochodów przekrywają się.



Rysunek 4.3: Empiryczne dopełnienia dystrybuant, $\Pi(m)$, rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej, m , z lat 2009–2010. Punkty czerwone reprezentują dane empiryczne pochodzące z badania EU-SILC a niebieskie dane empiryczne pochodzące z magazynu Forbes. Widać (podobnie jak dla lat 2004–2008), że dane ze wspólnego obszaru dochodów przekrywają się, jak być powinno. Rok 2009 zawiera wyjątkowo mało danych z magazynu Forbes – fakt ten jest omówiony w tekście.

Analizując powyższe wykresy możemy zauważyć, że dla roku 2009 istnieją tylko 3 obserwacje dotyczące gospodarstw domowych pochodzących z bazy danych magazynu Forbes. Innymi słowy, w roku tym nie istnieją gospodarstwa domowe uzyskujące wysokie dochody, czyli obserwuje się całkowity zanik klasy społecznej reprezentowanej przez te gospodarstwa. Ten niezwykle ważny fakt zostanie szerzej omówiony w rozdziale 4.4.

4.4 ANALIZA DANYCH EMPIRYCZNYCH

W następnych rozdziałach porównałem empiryczne dopełnienia dystrybuant, dotyczące rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej, z przewidywaniami modeli teoretycznych, tj. z prawem Boltzmanna–Gibbsa oraz słabym prawem Pareto. **Dzięki wyznaczeniu parametrów tych modeli, zweryfikowałem również zaproponowany w niniejszej rozprawie doktorskiej, rozszerzony model Yakovenko. Wykazałem przy tym, jak ważną rolę w analizie danych o dochodach gospodarstw domowych, odgrywa wykorzystanie danych empirycznych zawierających informacje o dochodach wszystkich trzech klas społecznych. Innymi słowy zauważyłem, że uwzględnienie w badaniach gospodarstw domowych o wysokich dochodach (poprzez odpowiednie połączenie danych EU–SILC i magazynu Forbes) ma istotny wpływ na otrzymane wyniki i czyni przedstawioną analizę pełniejszą.**

Podobnie, jak w przypadku gospodarstw domowych w Polsce (patrz rozdział 3.3), wyznaczyłem przy pomocy analizy rankingu dochodów wartość wykładnika Pareto dla gospodarstw domowych uzyskujących wysokie dochody.

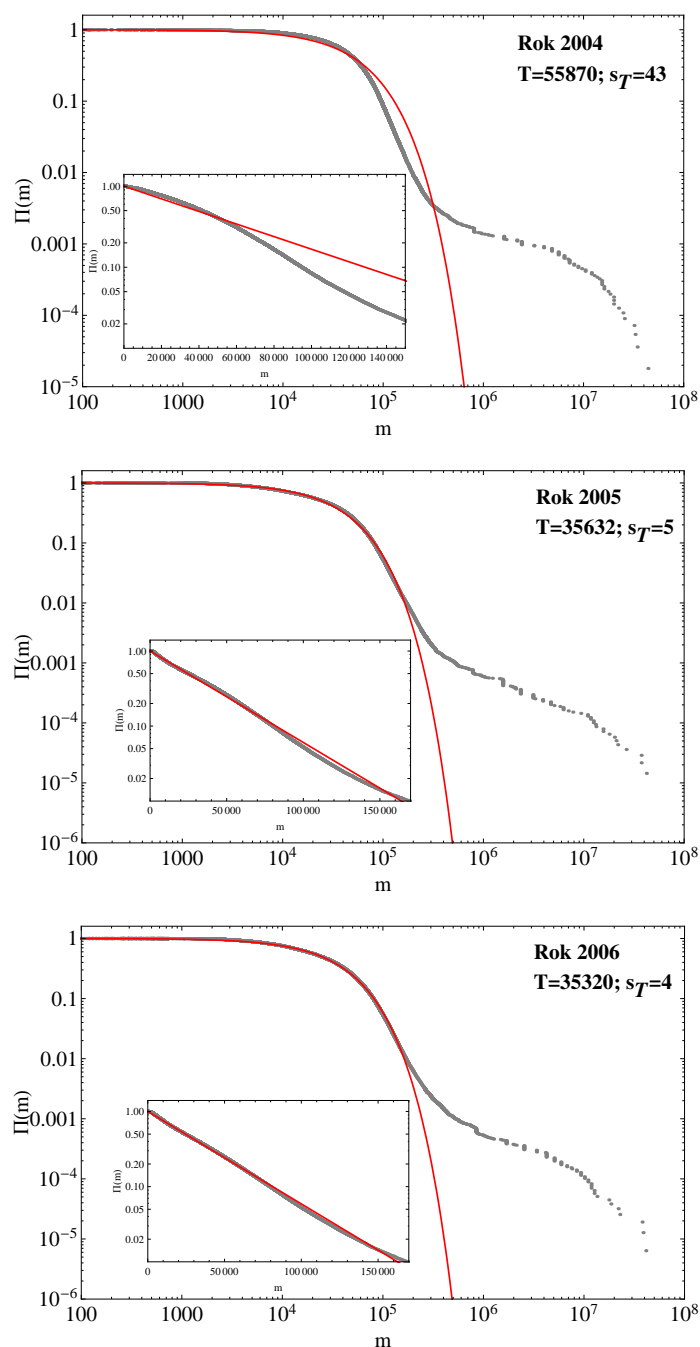
Zastosowanie prawa Boltzmanna–Gibbsa

W pierwszej kolejności do empirycznych dopełnień dystrybuant dotyczących rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej dopasowano dopełnienie dystrybuanty rozkładu wykładniczego wynikające z prawa Boltzmanna–Gibbsa:

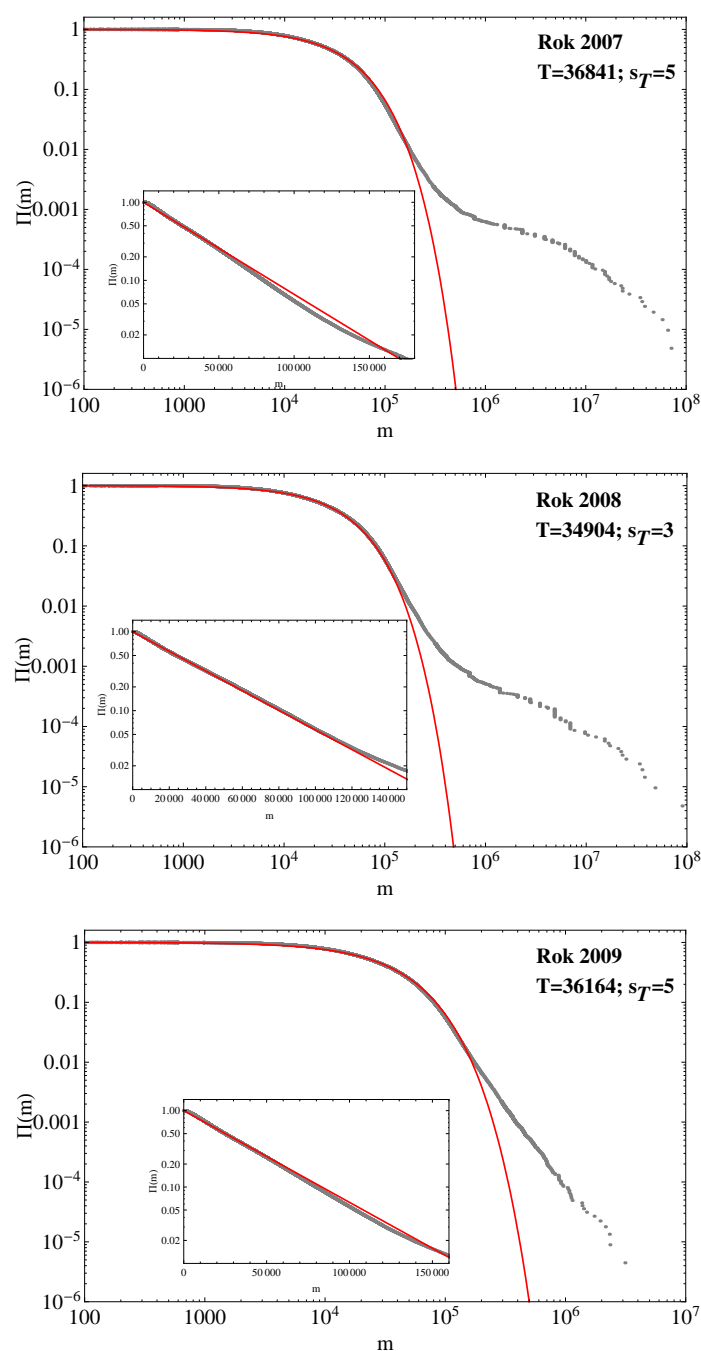
$$\Pi(m) = \exp\left(-\frac{m}{T}\right). \quad (4.1)$$

Wyniki w skali logarytmicznej zostały przedstawione na rysunkach 4.4 – 4.6⁵.

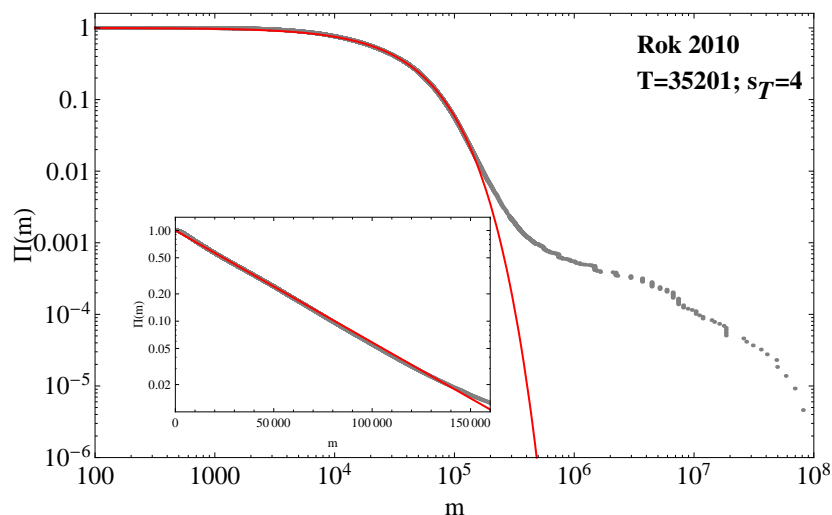
⁵ Występujące tutaj i na dalszych rysunkach oznaczenie s_x jest odchyleniem standardowym zmiennej x . W celu dopasowania krzywych posłużono się algorytmem Levenberga–Marquardta [98]. Dopełnienie dystrybuanty rozkładu wykładniczego dopasowano do wszystkich danych empirycznych dotyczących dochodów gospodarstw domowych w Unii Europejskiej.



Rysunek 4.4: Dopasowanie dopełnienia dystrybuanty rozkładu wykładniczego (linia ciągła) do empirycznego dopełnienia dystrybuanty rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej (punkty) z lat 2004–2006. Można zauważyć, że dla roku 2005 i 2006 prawo Boltzmann–Gibbsa dobrze opisuje gospodarstwa domowe o niskich dochodach. Na miniaturkach znajdują się te same wykresy w skali półlogarytmicznej w zakresach, dla których sprawdzamy czy dane empiryczne są wykładnicze.



Rysunek 4.5: Dopasowanie dopełnienia dystrybuanty rozkładu wykładniczego (linia ciągła) do empirycznego dopełnienia dystrybuanty rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej (punkty) z lat 2007–2009. Można zauważyć, że prawo Boltzmann–Gibbsa dobrze opisuje gospodarstwa domowe o niskich dochodach. Na miniaturkach znajdują się te same wykresy w skali półlogarytmicznej w zakresach, dla których sprawdzamy czy dane empiryczne są wykładnicze.



Rysunek 4.6: Dopasowanie dopełnienia dystrybuanty rozkładu wykładniczego (linia ciągła) do empirycznego dopełnienia dystrybuanty rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej (punkty) z roku 2010. Można zauważyć, że prawo Boltzmanna–Gibbsa dobrze opisuje gospodarstwa domowe o niskich dochodach, podobnie jak dla lat 2005–2009. Na miniaturce znajduje się ten sam wykres w skali półlogarytmicznej w zakresach, dla których sprawdzamy czy dane empiryczne są wykładnicze.

Prawo Boltzmanna–Gibbsa dobrze opisuje gospodarstwa domowe o niskich dochodach w latach 2005–2010. Potwierdza to jego uniwersalność w ekonomii, gdyż prawo to nadaje się również do opisu rocznych dochodów brutto (ang. *adjusted gross income*) osób w Stanach Zjednoczonych [2, 37, 45, 46, 50, 51].

Analizując powyższe wykresy możemy zauważyć, że w przypadku gospodarstw domowych o niskich dochodach empiryczne dopełnienia dystrybuant są nadzwyczaj stabilne i ich kształt nie zmienia się w kolejnych latach. Obserwacja ta prowadzi do wniosku, że uboga klasa społeczna znajduje się, w sensie fizyki statystycznej, w równowadze „termodynamicznej” [2], tj. w zasadzie z roku na rok nie ulega zmianie (rozpiętość temperatury wynosi tutaj zaledwie 5%; patrz tabela 4.3).

Wartości parametru T (ang. *income temperature*) dla poszczególnych dopasowań zostały zebrane w tabeli 4.3⁶.

⁶ Błędy parametrów T zostały zamieszczone w legendach na rysunkach 4.4 – 4.6. Z kolei, średnią $\langle x \rangle$ oraz rozpiętość R_x można oszacować na podstawie uzyskanych wartości parametru T .

Tabela 4.3: Parametry T uzyskane z dopasowania dopełnienia dystrybuanty rozkładu wykładniczego do empirycznego dopełnienia dystrybuanty rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej z lat 2005–2010. W przedostatnim wierszu znajduje się wartość średnia parametru T . Możemy zauważyć, że jego rozrzut względny (rozpiętość względna) R_x wynosi 5%.

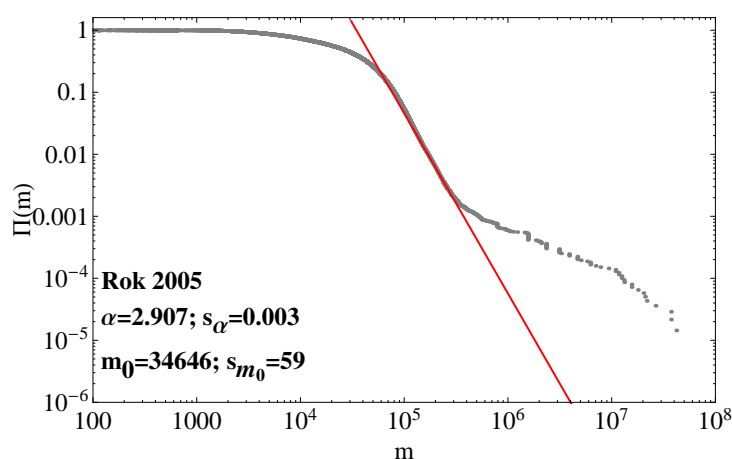
Rok	T
2005	35 632
2006	35 320
2007	36 841
2008	34 904
2009	36 164
2010	35 201
$\langle x \rangle$	35 677
$R_x = \frac{x_{\max} - x_{\min}}{\langle x \rangle}$	0.05

W przypadku roku 2004 prawo Boltzmanna–Gibbsa nie nadaje się do opisu gospodarstw domowych uzyskujących niskie dochody. Jest to prawdopodobnie spowodowane niekompletnością danych zbieranych na potrzeby badania EU–SILC. Badanie to zostało rozpoczęte w roku 2004 i wówczas tylko kilka krajów dostarczyło dane o dochodach (patrz tabela 4.1 oraz [104]). Dlatego też w dalszych analizach uwzględnione zostaną jedynie lata 2005–2010. Przypomnijmy, że w przypadku Polski gospodarstwa domowe o niskich dochodach opisywane są rozkładem log-normalnym, co w sposób zasadniczy odróżnia ją od Unii Europejskiej.

Zastosowanie słabego prawa Pareto

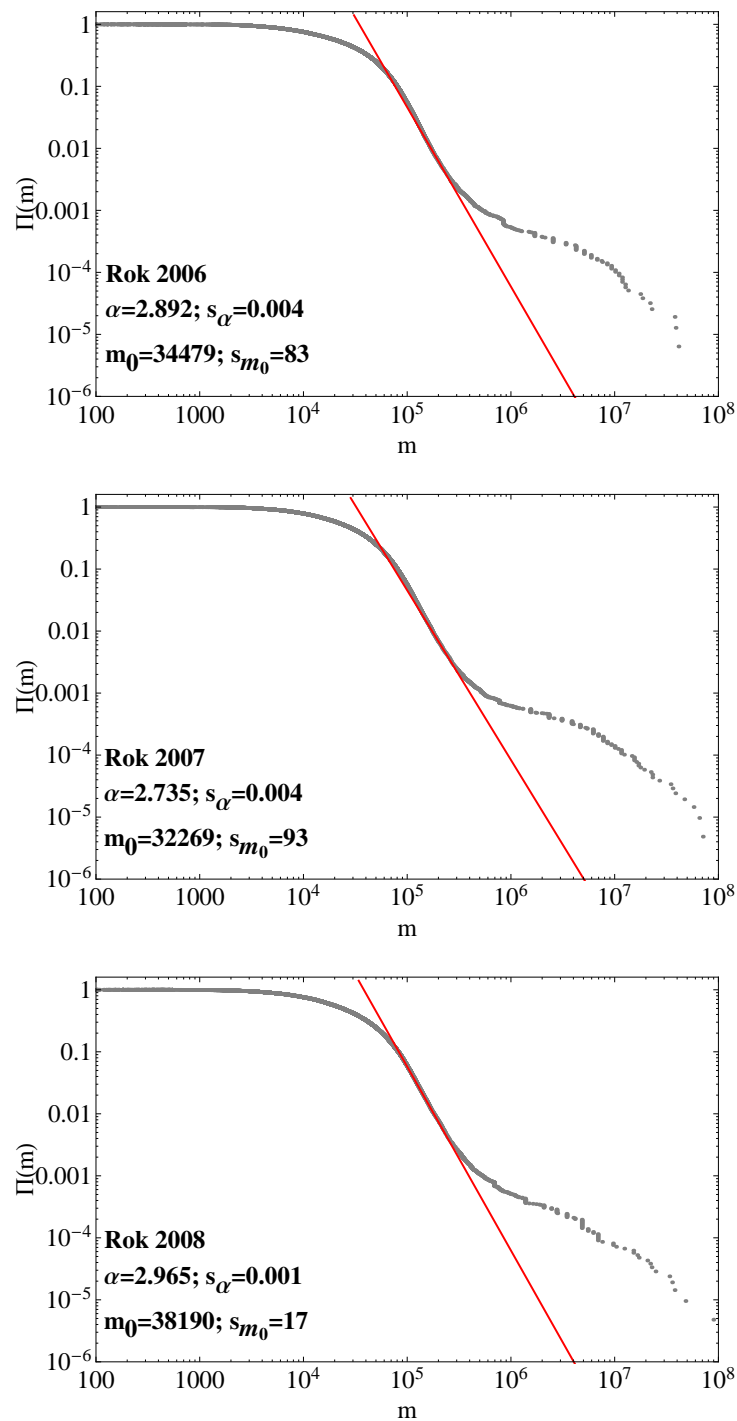
Do środkowych części omawianych empirycznych dopełnień dystrybuant dopasowano (patrz rys. 4.7 – 4.9) również słabe prawo Pareto⁷

$$\Pi(m) \sim (m/m_0)^{-\alpha}. \quad (4.2)$$

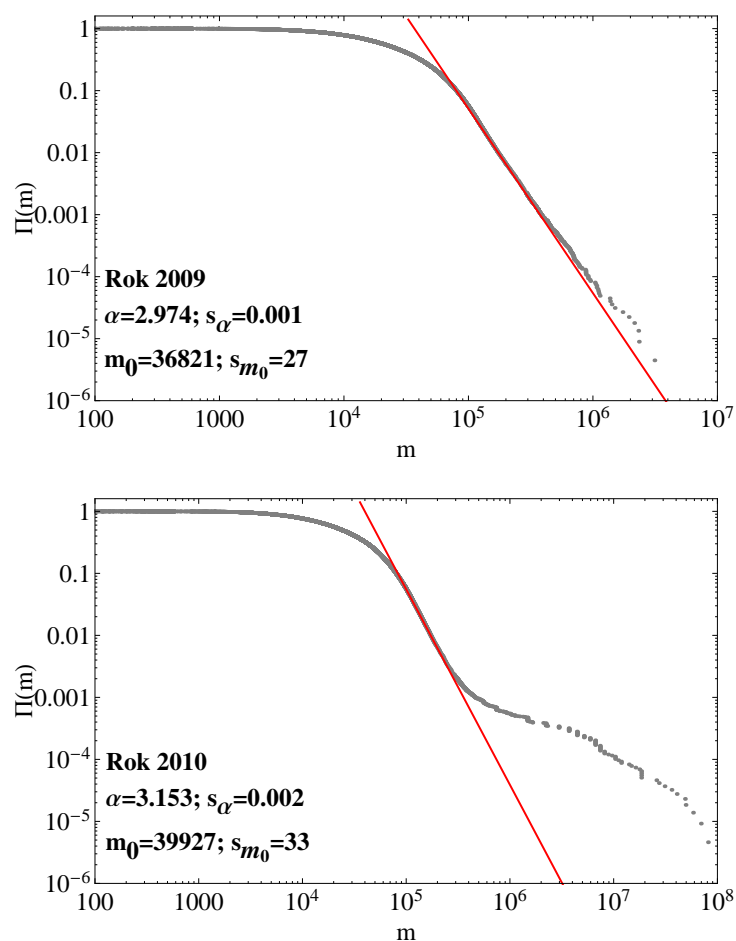


Rysunek 4.7: Dopasowanie słabego prawa Pareto (linia ciągła) do empirycznego dopełnienia dystrybuanty rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej (punkty) z roku 2005. Jak widać, słabe prawo Pareto dobrze opisuje gospodarstwa domowe o średnich dochodach.

⁷ W celu dopasowania krzywych posłużono się algorytmem Levenberga–Marquardta [98]. Dopełnienie dystrybuanty słabego prawa Pareto dopasowano do danych empirycznych dotyczących średnich dochodów gospodarstw domowych w Unii Europejskiej. Wyboru próbki dokonano tak, aby do środkowych części analizowanych rozkładów można było dopasować prostą w skali log-log. W tym celu należało znaleźć odpowiedni przedział dochodów. Przy wyborze próbki, do której fitowano prostą (w skali log-log) kierowano się maksymalizacją wartości R^2 .



Rysunek 4.8: Dopasowanie słabego prawa Pareto (linia ciągła) do empirycznego dopełnienia dystrybuanty rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej (punkty) z lat 2006–2008. Słabe prawo Pareto dobrze opisuje gospodarstwa domowe o średnich dochodach.



Rysunek 4.9: Dopasowanie słabego prawa Pareto (linia ciągła) do empirycznego dopełnienia dystrybuanty rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej (punkty) z lat 2009–2010. Słabe prawo Pareto dobrze opisuje gospodarstwa domowe o średnich dochodach. Raz jeszcze widzimy, że dopełnienie dystrybuanty dochodów z roku 2009 znacząco różni się od pozostałych.

Słabe prawo Pareto dobrze opisuje gospodarstwa domowe o średnich dochodach. Wartości wykładników Pareto dla poszczególnych lat zebrano w tabeli 4.4⁸.

⁸ Błędy wykładników zostały zamieszczone w legendach na rysunku 4.7 – 4.9. Z kolei, wartość średnią $\langle x \rangle$ oraz rozpiętość R_x można oszacować na podstawie uzyskanych wartości wykładników α .

Tabela 4.4: Wartości wykładników Pareto uzyskane z dopasowania słabego prawa Pareto do środkowej części empirycznego dopełnienia dystrybucyjności rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej z lat 2005–2010. W przedostatnim wierszu znajduje się wartość średnia wykładnika Pareto. Możemy zauważyć, że jego rozrzut względny (względna rozpiętość) R_x wynosi 14%.

Rok	α
2005	2.907
2006	2.892
2007	2.735
2008	2.965
2009	2.974
2010	3.153
$\langle x \rangle$	2.938
$R_x = \frac{x_{\max} - x_{\min}}{\langle x \rangle}$	0.14

W przypadku gospodarstw domowych uzyskujących średnie dochody wykładnik Pareto ulega niewielkim zmianom w czasie. Oznacza to, że poziom nierówności społecznych w ramach średniozamożnej klasy społecznej jest praktycznie stały. Niemniej, możemy zauważyć pewne tendencje. W latach 2005–2007 wykładnik Pareto malał, a zatem rosło rozwarstwienie społeczne wśród gospodarstw domowych o średnich dochodach. Z kolei, po roku 2007 (początek światowego kryzysu) obserwujemy wzrost wykładnika Pareto, a co zatem idzie zmniejszenie nierówności społecznych.

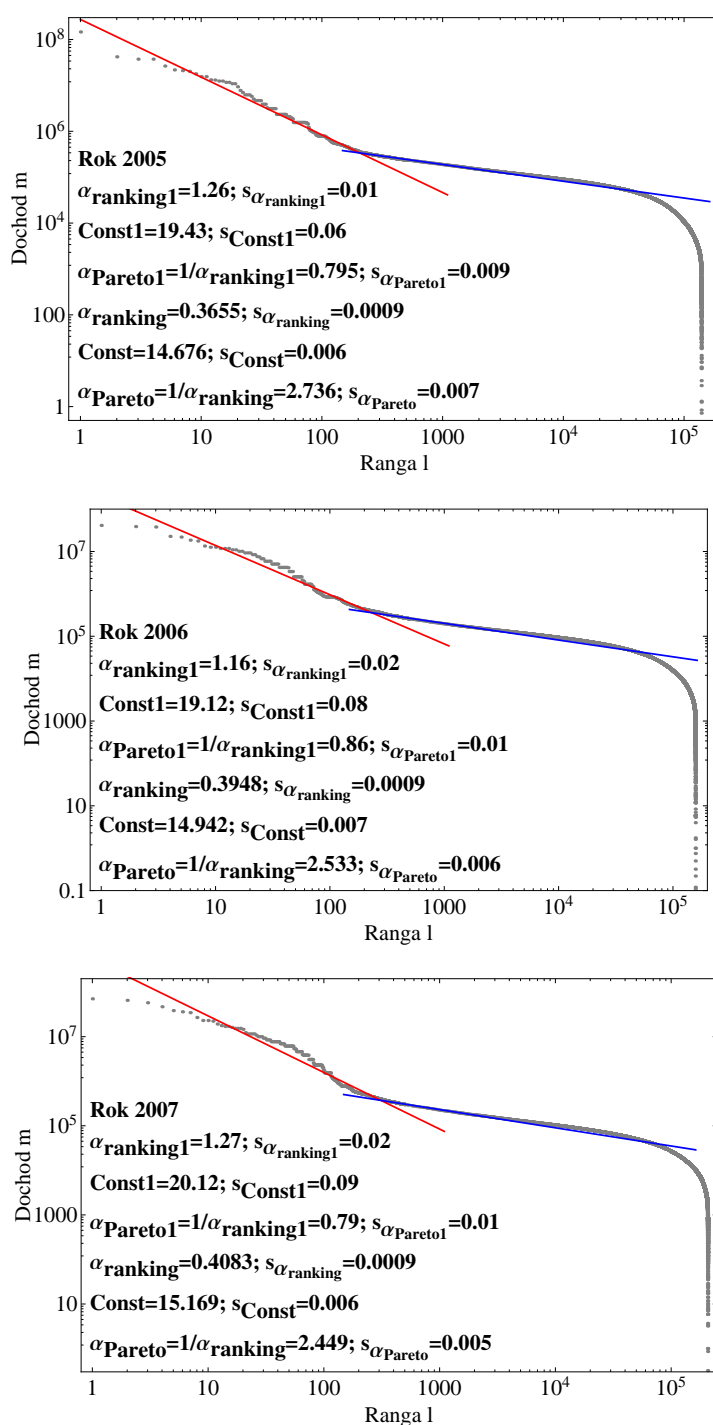
Podobnie jak dla polskich gospodarstw domowych o średnich dochodach, słabe prawo Pareto jest odpowiednie do zobrazowania europejskich gospodarstw domowych o średnich dochodach, ponieważ w tym wypadku zaczyna już odgrywać rolę konkurencja rynkowa (patrz rozdział 3.3).

Rankingi gospodarstw domowych o średnich i wysokich dochodach

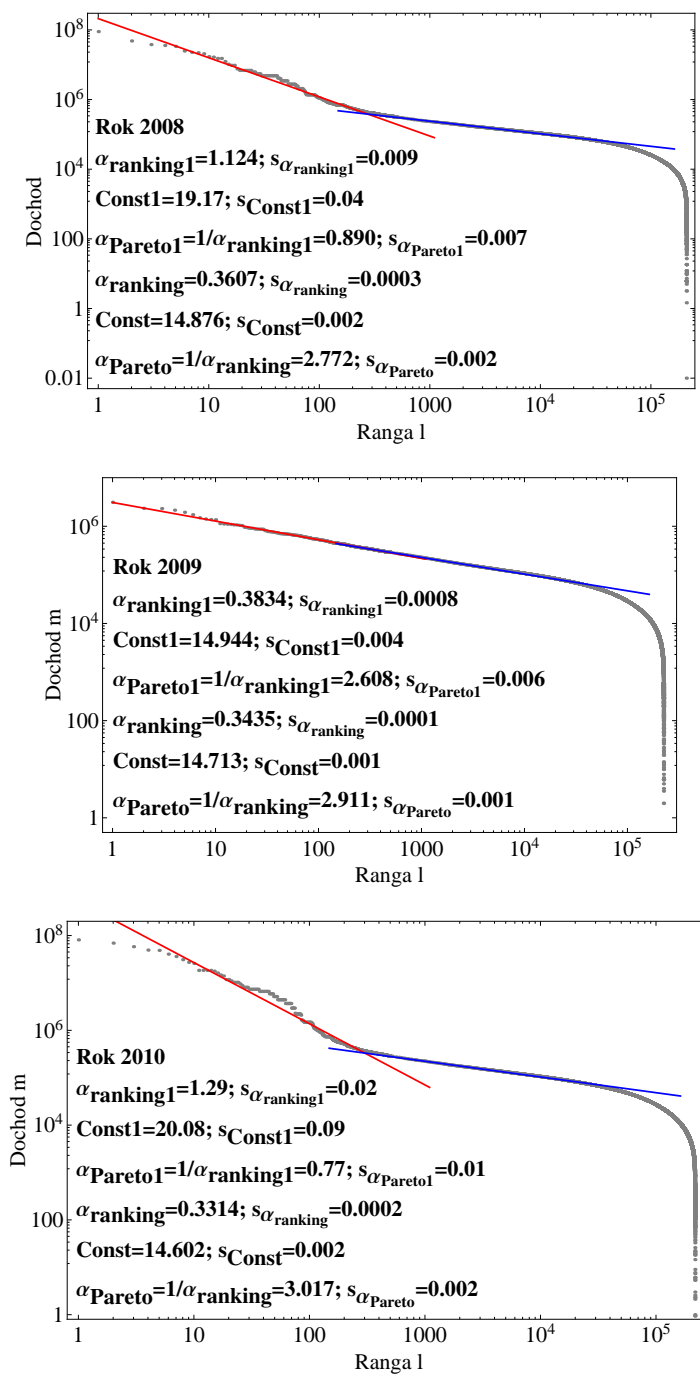
W kolejnym kroku, poprzez analizę rankingu gospodarstw domowych, wyznaczono wykładnik Pareto dla gospodarstw domowych uzyskujących wysokie dochody (metoda ta została opisana w rozdziale 3.3). Wyznaczenie wykładnika Pareto polegające na dopasowaniu prostej do „ogonów” empirycznych dopełnień dystrybuant w skali logarytmicznej jest również możliwe. Jednak sposób ten w przypadku danych dotyczących Unii Europejskiej wymaga przeskalowania danych (pomnożenia współrzędnych punktów przez czynnik rzędu 10^3). W przeciwnym razie, nie jesteśmy w stanie otrzymać wartości wykładnika Pareto ze względów numerycznych. Przeskalowanie danych uniemożliwia poprawne oszacowanie parametru m_0 pojawiającego się w słabym prawie Pareto. Zatem, nie jest również możliwe wykonanie wykresu słabego prawa Pareto dla oszacowanych parametrów. Mając na uwadze powyższe trudności, przy analizie wysokich dochodów gospodarstw domowych wykorzystano rankingi – wykresy w skali logarytmicznej znajdują się na rysunkach 4.10 – 4.11⁹.

Dodatkowo na rysunkach tych zamieszczono dopasowanie prostej dla gospodarstw domowych o średnich dochodach, aby jeszcze raz potwierdzić, że analiza rankingów stanowi podejście równoważne (ale numerycznie alternatywne) w stosunku do analizy dopełnień dystrybuant.

⁹ Na rysunkach 4.10 – 4.11 wielkości $Const$ oraz $Const1$ są stałymi pochodzącymi z dopasowania prostych (w skali log-log). Dopasowując do danych empirycznych (dotyczących średnich dochodów gospodarstw domowych w Unii Europejskiej) prostą w skali log-log, wyboru podpróbki dokonano analogicznie jak w przypadku słabego prawa Pareto. Natomiast, w przypadku danych o wysokich dochodach gospodarstw domowych w Unii Europejskiej, wyboru podpróbki dokonano w oparciu o pierwszą prostoliniową część rankingu (w skali log-log). W tym celu znaleziono graniczny dochód (odpowiadający określone miejsce w rankingu), poniżej którego dane są odrzucane. Zatem dopasowanie przeprowadzono w zakresie danych – od pewnego granicznego dochodu do dochodu najbogatszego gospodarstwa domowego dostępnego w próbie. Przy wyborze próbek, do których fitowano proste (w skali log-log) kierowano się maksymalizacją wartości R^2 .



Rysunek 4.10: Ranking rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej – lata 2005–2007. Czerwona linia stanowi dopasowanie prostej do wysokich dochodów gospodarstw domowych (parametry $\alpha_{\text{ranking}1}$, $\text{Const}1$, $\alpha_{\text{Pareto}1}$) a niebieska linia stanowi dopasowanie prostej do średnich dochodów gospodarstw domowych (parametry α_{ranking} , Const , α_{Pareto}).



Rysunek 4.11: Ranking rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej – lata 2008–2010. Czerwona linia stanowi dopasowanie prostej do wysokich dochodów gospodarstw domowych (parametry α_{ranking1} , Const1 , α_{pareto1}) a niebieska linia stanowi dopasowanie prostej do średnich dochodów gospodarstw domowych (parametry α_{ranking} , Const , α_{pareto}).

Gospodarstwa domowe o wysokich dochodach opisywane są w skali logarytmicznej prostą o nachyleniu $-\alpha_{\text{ranking1}}$. Zatem zgodnie z rozumowaniem przedstawionym w rozdziale 3.3, wykładnik Pareto możemy wyznaczyć z zależności $\alpha_{\text{Pareto1}} = \frac{1}{\alpha_{\text{ranking1}}}$. Wartości wykładników Pareto (dla gospodarstw domowych o wysokich dochodach) pochodzące z poszczególnych dopasowań zostały zebrane w tabeli 4.5. W tabeli tej umieszczono dla porównania wykładniki Pareto uzyskane z dopasowania słabego prawa Pareto do „ogonów” empirycznych dopełnień dystrybuant dotyczących rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej.

Tabela 4.5: Porównanie wykładników Pareto (gospodarstwa domowe o wysokich dochodach) uzyskanych z analizy rankingów dochodów gospodarstw domowych z wykładnikami Pareto uzyskanymi z dopasowania słabego prawa Pareto do empirycznego dopełnienia dystrybuanty rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej z lat 2005–2010. W przedostatnim wierszu znajdują się wartości średnie parametrów. Możemy zauważyć, że rozrzut względny (względna rozpiętość) R_x wartości wykładników Pareto jest znaczący. Wynika to, że skokowego wzrostu wykładnika Pareto w roku 2009.

Rok	α_{Pareto1} (ranking)	α_{Pareto1} (dopełnienie dystrybuanty)
2005	0.795 ± 0.009	0.736 ± 0.009
2006	0.86 ± 0.01	0.80 ± 0.01
2007	0.79 ± 0.01	0.635 ± 0.009
2008	0.890 ± 0.007	0.95 ± 0.01
2009	2.608 ± 0.006	2.728 ± 0.006
2010	0.77 ± 0.01	0.652 ± 0.008
$\langle x \rangle$	1.119 ± 0.009	1.083 ± 0.004
$R_x = \frac{x_{\text{max}} - x_{\text{min}}}{\langle x \rangle}$	1.64 ± 0.02	1.93 ± 0.01

Otrzymane wartości wykładnika Pareto są bliskie jedności. Jest to zgodne z podobnymi analizami dotyczącymi wysokich dochodów gospodarstw domowych w innych krajach [2, 3, 5, 30, 44, 67, 102, 103]. Zatem można oczekiwać, że mechanizm zdobywania wysokich dochodów przez gospodarstwa domowe Unii Europejskiej,

opiera się na istnieniu konkurencji rynkowej (podobnie jak dla gospodarstw domowych o średnich dochodach) [100], jednak w tym przypadku jednostki tworzące gospodarstwa domowe są przeważnie właścicielami konkurujących firm, przy czym dochody tych firm opisywane są prawem Zipfa¹⁰ [88,91,92]. Model mikroskopowy odtwarzający słabe prawo Pareto z wykładnikiem $\alpha = 1$, a więc zbliżonym do uzyskanego tutaj został omówiony w rozdziale 2.1.

Można zauważyć, że w roku 2009 wartość wykładnika Pareto jest znacząco różna od jedności. Rok ten stanowi przypadek szczególny ze względu na zniknięcie klasy społecznej reprezentowanej przez gospodarstwa domowe o wysokich dochodach. Fakt ten zostanie wyjaśniony i zinterpretowany w dalszej części niniejszego rozdziału.

Pozostałe uzyskane wartości wykładnika Pareto są mniejsze od jedności. W przypadku kiedy $\alpha_{\text{pareto}} \leq 1$, momenty (pierwszy i wyższych rzędów) gęstości prawdopodobieństwa opisującej słabe prawo Pareto są rozbieżne. Zatem, jak wspomniano w rozdziale 1.4, w takim przypadku należy używać kwantyli (które są zawsze skończone [87]).

Na podstawie rysunków 4.10 – 4.11 możemy również stwierdzić, że gospodarstwa domowe o średnich dochodach opisywane są w skali logarytmicznej prostą o nachyleniu $-\alpha_{\text{ranking}}$. Jak należało oczekiwać, odwrotność α_{ranking} jest tutaj bardzo zbliżona do wyznaczonych wartości wykładnika Pareto, α_{pareto} , uzyskanych w przypadku dopełnień dystrybuant rocznych dochodów brutto gospodarstw domowych. Wskazuje to na zgodność obu zastosowanych metod [90]. Porównanie obu wykładników zamieszczono w tabeli 4.6¹¹.

¹⁰ Prawo Zipfa jest prawem Pareto z wykładnikiem $\alpha = 1$.

¹¹ Błędy wykładników zostały zamieszczone w legendach na rysunkach 4.7 – 4.9 oraz 4.10 – 4.11. Z kolei, wartość średnią $\langle x \rangle$ oraz rozpiętość R_x można oszacować na podstawie uzyskanych wartości wykładników α_{pareto} .

Tabela 4.6: Porównanie wykładników Pareto (dla gospodarstw domowych o średnich dochodach) uzyskanych z analizy rankingów dochodów gospodarstw domowych z wykładnikami Pareto uzyskanymi z dopasowania słabego prawa Pareto do empirycznego dopełnienia dystrybuanty rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej w latach 2005–2010. W przedostatnim wierszu znajdują się wartości średnie wykładników Pareto. Możemy zauważyć, że rozrzut względny (względna rozpiętość) R_x wykładników Pareto nie przekracza 21%.

Rok	α_{Pareto} (ranking)	α_{Pareto} (dopełnienie dystrybuanty)
2005	2.736	2.907
2006	2.533	2.892
2007	2.449	2.735
2008	2.772	2.965
2009	2.911	2.974
2010	3.017	3.153
$\langle x \rangle$	2.736	2.938
$R_x = \frac{x_{\max} - x_{\min}}{\langle x \rangle}$	0.21	0.14

Weryfikacja rozszerzonego modelu Yakovenko

Obecnie dokonam weryfikacji zaproponowanego przeze mnie rozszerzonego modelu Yakovenko. W modelu tym uzyskujemy teoretyczne dopełnienie dystrybuanty opisujące dochody wszystkich trzech klas społecznych tj. gospodarstw domowych o niskich, średnich i wysokich dochodach, włączając w to obszary przejściowe pomiędzy tymi klasami. **Przedstawione wyniki analizy danych empirycznych stanowią jeden z kluczowych elementów niniejszej rozprawy doktorskiej.**

Rozszerzony model Yakovenko jest opisywany gęstością prawdopodobieństwa postaci (patrz rozdział 2.2 a tam wzór (2.37)):

$$P(m) = \begin{cases} C' \frac{\exp\left(-\frac{m_0}{T_1} \arctan \frac{m}{m_0}\right)}{\left[1 + \left(\frac{m}{m_0}\right)^2\right]^{\frac{\alpha+1}{2}}}, & \text{dla } m < m_1, \\ C'' \frac{\exp\left(-\frac{m_0}{T_1} \arctan \frac{m}{m_0}\right)}{\left[1 + \left(\frac{m}{m_0}\right)^2\right]^{\frac{\alpha_1+1}{2}}}, & \text{dla } m \geq m_1. \end{cases} \quad (4.3)$$

Dla gęstości prawdopodobieństwa opisanej równaniem (4.3) nie można wyznaczyć dopełnienia dystrybucyjnego $\Pi(m)$ w zamkniętej postaci analitycznej. Niemniej jednak, wartości tego rozkładu można obliczyć przy użyciu metod numerycznych.

Zasadniczym problemem przy podejściu numerycznym jest sposób, w jaki należy porównać tak skomplikowane teoretyczne dopełnienie dystrybucyjnego do danych empirycznych. Zaproponowana przeze mnie procedura porównania składa się z trzech etapów i polega na wyznaczeniu parametrów występujących w równaniu (4.3) na podstawie oszacowań parametrów uzyskanych z dopasowania do empirycznych dopełnień dystrybucyjnego prawa Boltzmanna–Gibbsa, słabego prawa Pareto oraz analizy rankingu gospodarstw domowych o wysokich dochodach.

W pierwszym kroku należy oszacować wartości parametrów m_0 i m_1 . Zgodnie z tym, co powiedziano w rozdziale 1.5 parametr m_0 można traktować jako granicę między prawem Boltzmanna–Gibbsa a słabym prawem Pareto. Zatem jego wartość można oszacować odczytując współrzędne punktu przecięcia dopełnienia dystrybucyjnego rozkładu wykładniczego (prawo Boltzmanna–Gibbsa) z dopełnieniem dystrybucyjnego, pochodzącym ze słabego prawa Pareto (dla gospodarstw domowych uzyskujących średnie dochody). Podobnie wyznacza się parametr m_1 , który traktujemy jako graniczną wartość dochodu pomiędzy gospodarstwami domowymi o średnich i wysokich dochodach (patrz rozdział 2.2). Jego wartość szacujemy, odczytując współrzędne punktu przecięcia dwóch dopełnień dystrybucyjnego, opisujących słabe prawo Pareto (odpowiednio dla gospodarstw domowych o średnich dochodach i gospodarstw domowych o wysokich dochodach)¹². W tym przypadku,

¹² Analizując rankingi gospodarstw domowych o wysokich dochodach zaznaczono, że wykonanie wykresu słabego prawa Pareto w przypadku tych gospodarstw nie jest możliwe, ze względu na niemożność poprawnego oszacowania parametru m_0 (parametr ten pochodzi ze słabego prawa Pareto i nie należy mylić go tutaj z parametrem m_0 występującym w rozszerzonym modelu Yakovenko; przypomnijmy, że w przypadku silnego

współrzedną tego punktu na osi odciętych należy taktować jako wartość wyjściową (przybliżoną) przy poszukiwaniu najlepszego oszacowania parametru m_1 . Z reguły wartość m_1 , jaką należy wstawić do rozszerzonego modelu Yakovenko (aby otrzymać poprawne dopasowanie) jest o około 100 000 Euro większa (jest jeszcze o tym mowa poniżej).

W kolejnym kroku wyznacza się parametry T i T_1 . Wartość parametru T pojawiającego się w równaniu (4.3) uzyskujemy poprzez dopasowanie prawa Boltzmann-Gibbsa do empirycznego dopełnienia dystrybuanty dotyczącego rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej (patrz rysunki 4.4 – 4.6). Z kolei „temperatura” T_1 wyraża średni dochód gospodarstw domowych (patrz rozdział 1.3), należących do grupy gospodarstw domowych o wysokich dochodach. Zatem zauważając, że ten średni dochód nie może być niższy, niż najniższy dochód gospodarstwa domowego z grupy gospodarstw domowych o wysokich dochodach (czyli dochodach większych, bądź równych m_1), wartość parametru T_1 przyjmujemy po prostu równą wartości parametru m_1 (co zmniejsza liczbę parametrów swobodnych modelu).

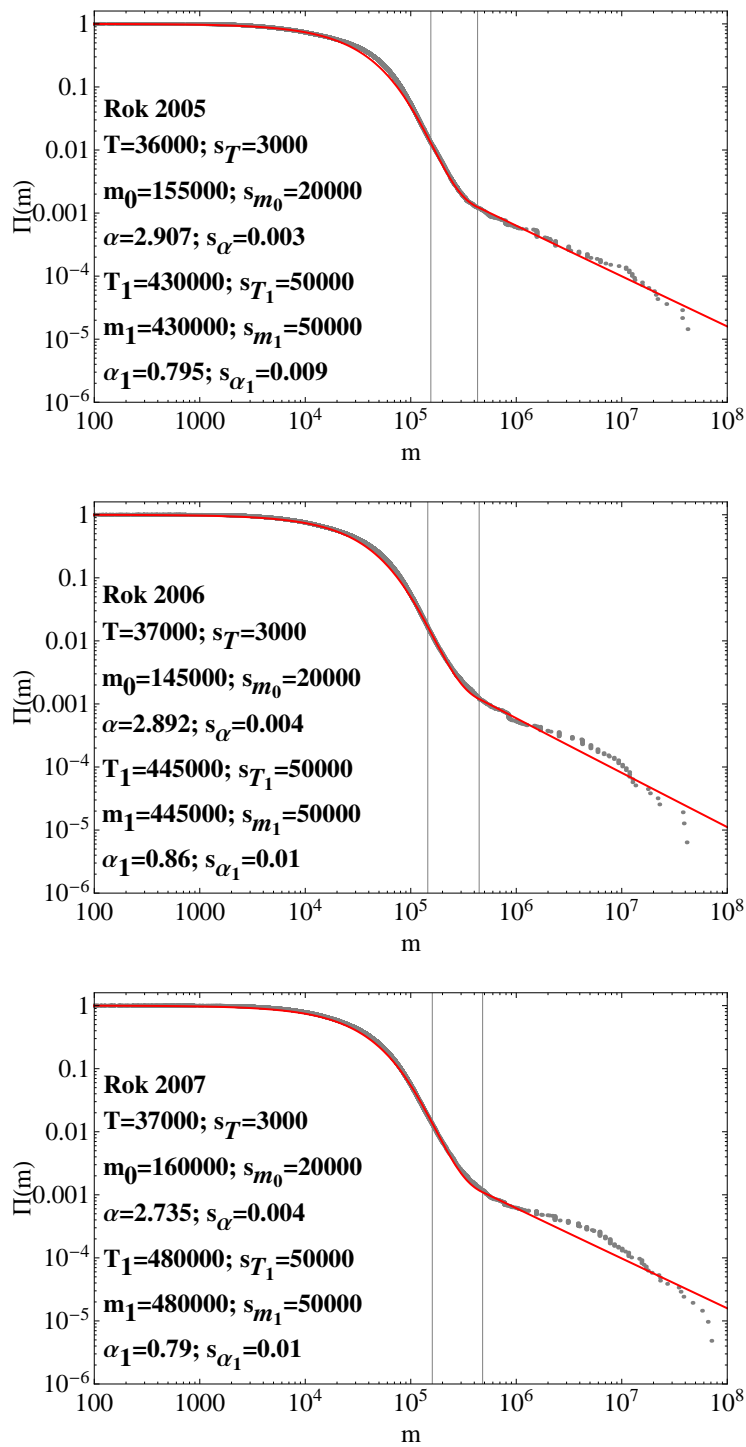
W ostatnim kroku szacujemy wartości parametrów α oraz α_1 . Parametr α uzyskujemy na podstawie dopasowania słabego prawa Pareto do środkowej części empirycznego dopełnienia dystrybuanty, odpowiadającej gospodarstwom domowym o średnich dochodach (rysunki 4.7 – 4.9). Z kolei, parametr α_1 przyjmuje wartość wykładnika Pareto (α_{Pareto1}) uzyskanego na podstawie analizy rankingu gospodarstw domowych o wysokich dochodach (rysunki 4.10 – 4.11).

Dysponując teraz wszystkimi parametrami pojawiającymi się w rozszerzonym modelu Yakovenko, porównano (na drodze numerycznej) teoretyczne dopełnienie dystrybuanty z empirycznym dopełnieniem dystrybuanty dotyczącym rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej. Możliwość tego

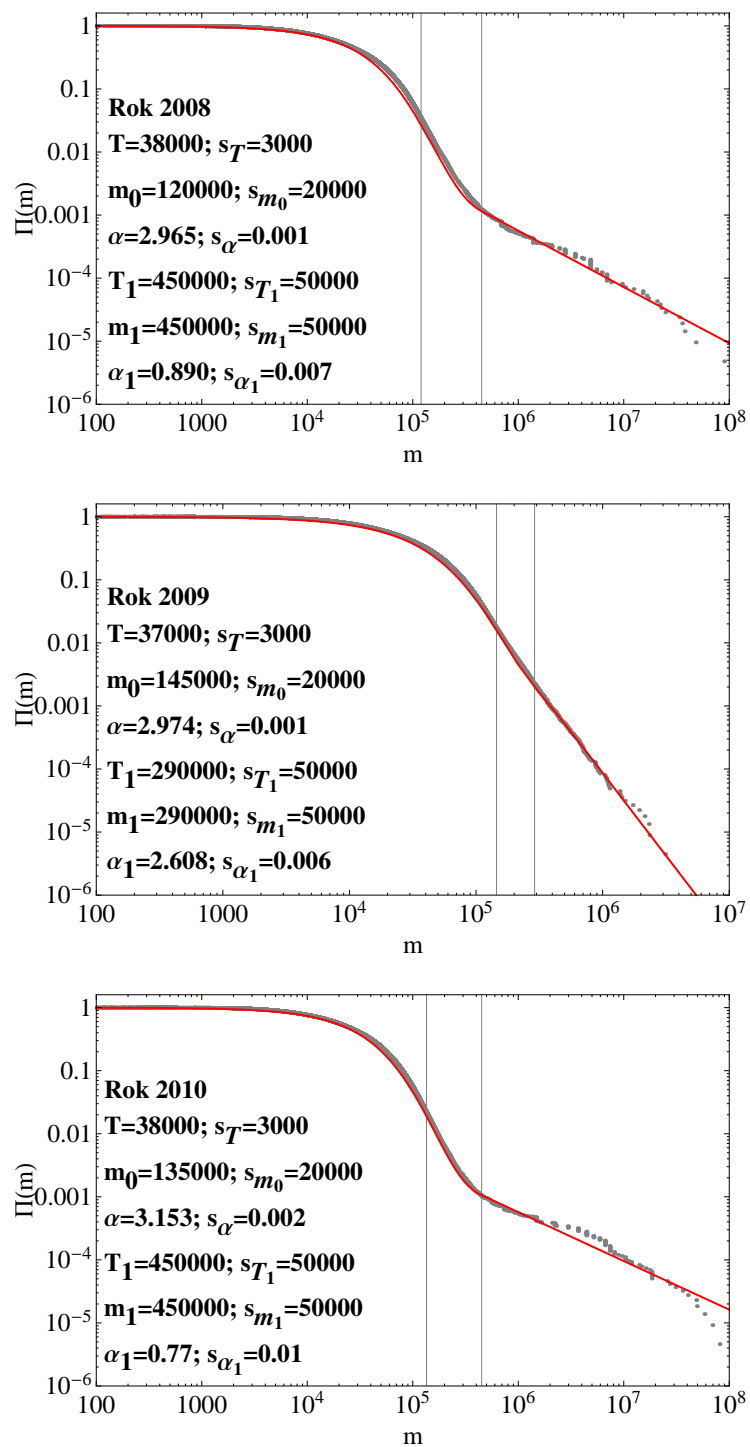
prawa Pareto parametr m_0 oznacza graniczną, najniższą wartość dochodu, jaką mogą osiągać gospodarstwa domowe, natomiast w słabym prawie Pareto jest on parametrem swobodnym i nie posiada interpretacji). Jednak wartość parametru m_0 można oszacować dobierając jego wartość tak, aby otrzymany wykres słabego prawa Pareto (dla wyznaczonego z rankingu wykładnika Pareto) pokrył się z „ogonem” empirycznego dopełnienia dystrybuanty. Procedura ta jest wystarczająca do wyznaczenia punktu przecięcia słabego prawa Pareto dla gospodarstw domowych o średnich dochodach ze słabym prawem Pareto dla gospodarstw domowych o wysokich dochodach.

typu porównania oznacza, że w rozszerzonym modelu Yakovenko nie pozostały już swobodne parametry, co dobrze świadczy o jego jakości. Wyniki porównania zostały zamieszczone na rysunkach 4.12 – 4.13 (skala logarytmiczna)¹³.

13 Aby otrzymać lepsze dopasowania, w przypadku lat 2008 i 2010 nieznacznie zwiększono wartość parametru T otrzymanego z dopasowania prawa Boltzmann–Gibbsa do empirycznego dopełnienia dystrybucji rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej z roku 2008 oraz 2010. Błędy parametrów T , T_1 , m_0 i m_1 zostały oszacowane poprzez zmienianie wartości tych parametrów i analizie poprawności dopasowania otrzymanych w ten sposób teoretycznych dopełnień dystrybucji do dopełnień empirycznych. Błąd wykładnika α został oszacowany poprzez dopasowanie do empirycznych dopełnień dystrybucji słabego prawa Pareto dla gospodarstw domowych o średnich dochodach (rysunki 4.7 – 4.9), natomiast błąd wykładnika α_1 oszacowano na podstawie analizy rankingów gospodarstw domowych o wysokich dochodach (rysunki 4.10 – 4.11).



Rysunek 4.12: Porównanie przewidywań rozszerzonego modelu Yakovenko (linia ciągła) z empirycznym dopełnieniem dystrybuanty rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej (punkty) z lat 2005–2007. Cienkie linie pionowe zaznaczają odpowiednio wartości m_0 oraz m_1 .



Rysunek 4.13: Porównanie przewidywań rozszerzonego modelu Yakovenko (linia ciągła) z empirycznym dopełnieniem dystrybuanty rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej (punkty) z lat 2008–2010. Cienkie linie pionowe zaznaczają odpowiednio wartości m_0 oraz m_1 .

Jak widać, rozszerzony model Yakovenko dobrze opisuje empiryczne dopełnienia dystrybuant dotyczące rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej. Zatem, przy pomocy tego modelu możemy opisać dochody wszystkich trzech klas społecznych, tj. gospodarstw domowych o niskich, średnich i wysokich dochodach. Tak dobra zgodność z danymi empirycznymi została uzyskana dzięki przyjęciu przede wszystkim następujących założeń:

- rozszerzony model Yakovenko dopuszcza współistnienie procesu addytywnego i multiplikatywnego przy jednoczesnym zróżnicowaniu dynamiki zmian dochodu – w modelu tym przyjmuje się, że gospodarstwa domowe o niskich i średnich dochodach bogacą się (bądź tracą dochody) inaczej niż gospodarstwa domowe o wysokich dochodach (patrz rozdział 2.2; równanie (2.25) oraz (2.26)),
- rozszerzony model Yakovenko spełnia warunek ciągłości gęstości prawdopodobieństwa (rozdział 2.2; równanie (2.27)).

Wartości parametrów dla poszczególnych porównań rozszerzonego modelu Yakovenko zamieszczono w tabeli 4.7¹⁴.

¹⁴ Błędy parametrów zostały zamieszczone w legendach na rysunkach 4.12 – 4.13. Z kolei, wartość średnią $\langle x \rangle$ oraz rozpiętość R_x można oszacować na podstawie uzyskanych parametrów.

Tabela 4.7: Parametry uzyskane z porównania rozszerzonego modelu Yakovenko z empirycznym dopełnieniem dystrybuanty rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej z lat 2005–2010. W przedostatnim wierszu znajdują się wartości średnie parametrów. Możemy zauważyć, że rozrzut względny parametrów T_1 , m_1 oraz α_1 jest znaczący. Wynika to, że skokowego wzrostu wykładnika Pareto α_1 w roku 2009, przy jednoczesnym skokowym spadku wartości parametrów T_1 oraz m_1 .

Rok	T	m_0 [EUR]	α	T_1	m_1 [EUR]	α_1
2005	36 000	155 000	2.907	430 000	430 000	0.795
2006	37 000	145 000	2.892	445 000	445 000	0.86
2007	37 000	160 000	2.735	480 000	480 000	0.79
2008	38 000	120 000	2.965	450 000	450 000	0.890
2009	37 000	145 000	2.974	290 000	290 000	2.608
2010	38 000	135 000	3.153	450 000	450 000	0.77
$\langle x \rangle$	37 167	143 333	2.938	424 167	424 167	1.119
$R_x = \frac{x_{\max} - x_{\min}}{\langle x \rangle}$	0.05	0.28	0.14	0.45	0.45	1.64

Analizując wartości parametrów zamieszczone w tabeli 4.7 możemy stwierdzić, że parametr T nieznacznie rośnie w czasie, a zatem dla gospodarstw domowych o niskich dochodach rośnie średni dochód przypadający na pojedyncze gospodarstwo domowe, co jest głównie spowodowane inflacją, a nie bogaceniem się tej grupy społecznej.

Z kolei, parametr m_0 oscyluje wokół wartości 140 000 Euro. Wartość tę można uznać za graniczną wartość dochodu pomiędzy gospodarstwami domowymi o niskich dochodach a gospodarstwami o średnich dochodach. Podobnie jest dla parametru m_1 , który dla lat 2005–2008 oraz 2010 oscyluje wokół wartości 450 000 Euro, oznaczającej graniczną wartość dochodu pomiędzy gospodarstwami domowymi o średnich dochodach a gospodarstwami o wysokich dochodach.

Analiza zmian wartości parametru α (tabela 4.7) pozwala zauważyć, że dla lat 2005–2007 parametr ten ma tendencję malejącą, po czym po roku 2007 zaczyna rosnąć. Oznacza to, że w latach 2005–2007 rozwarstwienie społeczne, w ramach

klasy społecznej reprezentowanej przez gospodarstwa domowe o średnich dochodach, rosło a następnie w latach 2008–2010 – malało. Z kolei, parametr α_1 w latach 2005–2008 oraz 2010 zmieniał się stosunkowo nieznacznie. Innymi słowy, rozwarstwienie społeczne w ramach klasy społecznej reprezentowanej przez gospodarstwa domowe o wysokich dochodach pozostawało na niezmiennym, wysokim poziomie (gdyż $\alpha_1 < 1$).

W przypadku roku 2009 obserwujemy, że wartości parametrów T_1 , m_1 oraz α_1 różnią się znacząco względem pozostałych lat. W roku tym nastąpił skokowy spadek wysokich dochodów gospodarstw domowych, a także skokowy spadek samej liczby gospodarstw domowych uzyskujących wysokie dochody¹⁵. Jest to spowodowane kryzysem ekonomicznym zapoczątkowanym w 2007 roku, którego apogeum nastąpiło w roku 2008 (a zatem wpłynęło na dochody uzyskane w roku kolejnym, tj. 2009). Kryzys ten doprowadził do obniżenia granicznej wartości dochodu pomiędzy gospodarstwami domowymi uzyskującymi średnie dochody a gospodarstwami domowymi uzyskującymi wysokie dochody. Przyczynił się również do znaczącego zmniejszenia rozwarstwienia społecznego w ramach klasy społecznej reprezentowanej przez gospodarstwa domowe o wysokich dochodach (w roku 2009 nastąpił znaczący wzrost α_1 do wartości 2.608), czyniąc ją, praktycznie rzecz biorąc, klasą średnią. Możemy zauważyć, że w roku 2009 klasa społeczna reprezentowana przez gospodarstwa domowe o wysokich dochodach w zasadzie nie istnieje – tylko niewielka liczba gospodarstw domowych uzyskiwała w tym roku wysokie dochody. Pozostałe gospodarstwa domowe, które dotychczas zaliczały się do tej klasy społecznej, uzyskiwały dochody o wiele niższe a przez to zaczęły przynależeć do klasy społecznej reprezentowanej przez gospodarstwa domowe o średnich bądź niskich dochodach. Należy jednak pamiętać, że w latach poza kryzysem, kształt obserwowanych empirycznych dopełnień dystrybuant jest niezwykle stabilny. Obserwujemy jedynie zmianę liczby gospodarstw domowych należących do określonych klas społecznych, a struktura dochodów społeczeństwa (jako całości) w zasadzie nie ulega zmianom.

¹⁵ Dla roku 2009 w bazie danych magazynu Forbes mamy tylko 3 obserwacje.

4.5 DODATKOWE UWAGI

Poniżej przedstawię dodatkowe rozważania wynikające z analizy danych empirycznych przeprowadzonej w rozdziale 4.4.

Równanie Langevina a ryzyko rynkowe

Zgodność rozszerzonego modelu Yakovenko z danymi empirycznymi potwierdza możliwość opisu rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej przy pomocy dynamiki Langevina. Dynamika ta wyrażona jest poprzez nieliniowe, progowe równanie Langevina (równanie 1.2) ze współczynnikiem dryfu $A(m)$ określonym zależnością (2.25) oraz wariancją¹⁶ $B(m)$, zdefiniowaną równaniem (2.26). Dla wartości dochodu równej m_1 mamy do czynienia ze skokiem współczynnika proporcjonalności (występującego w wyrażeniu $A(m)$), który zmienia się z a na a' , przy czym $a' < a$ (gdyż¹⁷: $\alpha_1 < \alpha$; patrz tabela 4.7). Oznacza to, że dla $m > m_1$ siła stochastyczna $C(m)\eta(t)$ w równaniu (1.2) (pamiętając, że: $B(m) = \frac{C^2(m)}{2}$) ma relatywnie większy wpływ na dynamikę Langevina niż siła oporu lepkiego $-A(m)$. Innymi słowy, jeśli wariancję $B(m)$, potraktujemy (jak to się zwykle robi) jako miarę ryzyka [89], to możemy powiedzieć, że działalność gospodarcza (w sensie aktywności ekonomicznej) gospodarstw domowych o wysokich dochodach jest obciążona wyraźnie większym ryzykiem niż działalność gospodarstw domowych o niskich i średnich dochodach.

Rozszerzony model Yakovenko a procedura łączenia baz danych

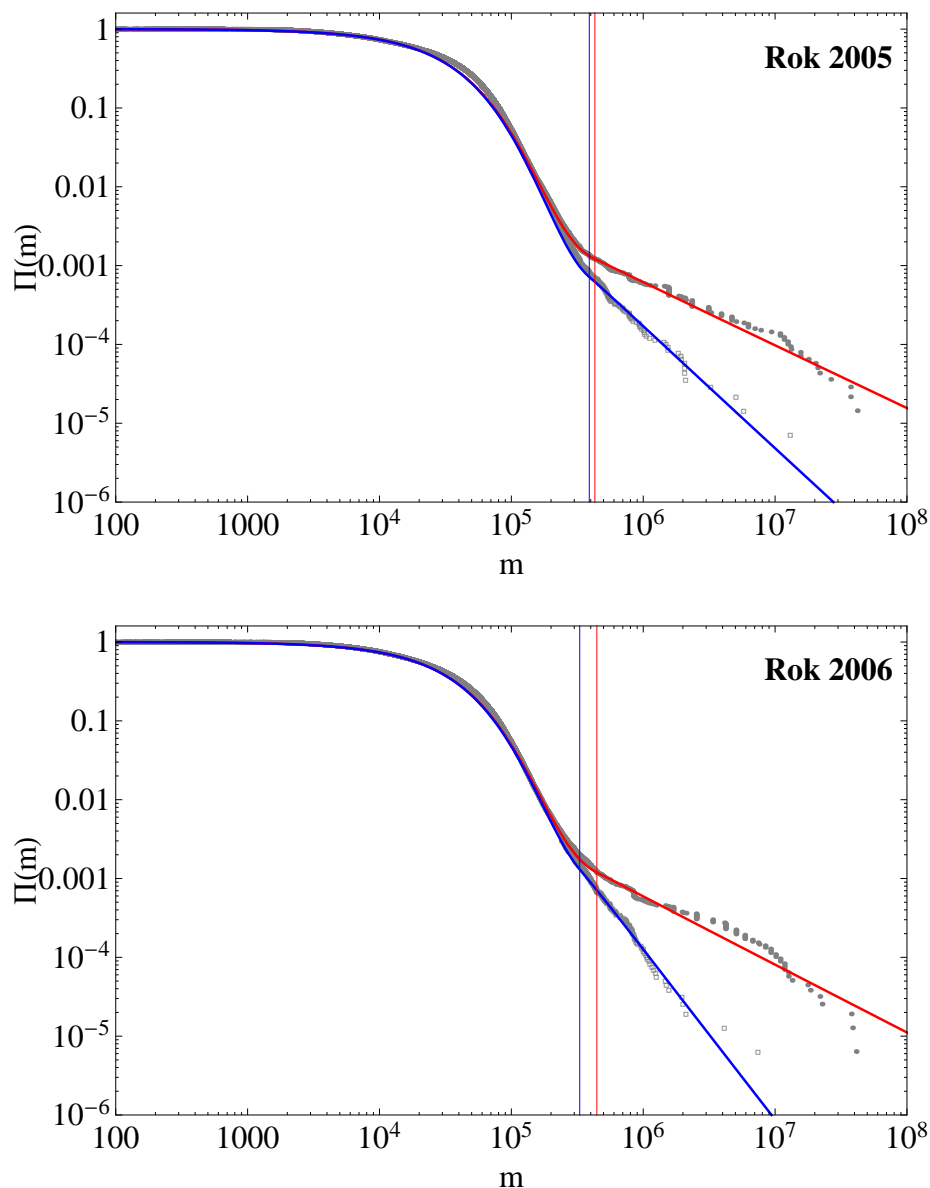
Należy zaznaczyć, że procedura łączenia baz danych przedstawiona w rozdziale 4.2 wpłynęła, zgodnie z oczekiwaniami, jedynie na zmianę struktury dochodów (a więc i kształt empirycznych dopełnień dystrybuant) w ramach klasy społecznej reprezentowanej przez gospodarstwa domowe o wysokich dochodach. Natomiast, nie miała ona wpływu na strukturę dochodów w przypadku klas społecznych reprezentowanych odpowiednio przez gospodarstwa domowe o niskich i średnich dochodach.

¹⁶ Dokładniej rzecz biorąc, jest to połowa wariancji (patrz równanie (1.28)), o ile wariancja ta istnieje.

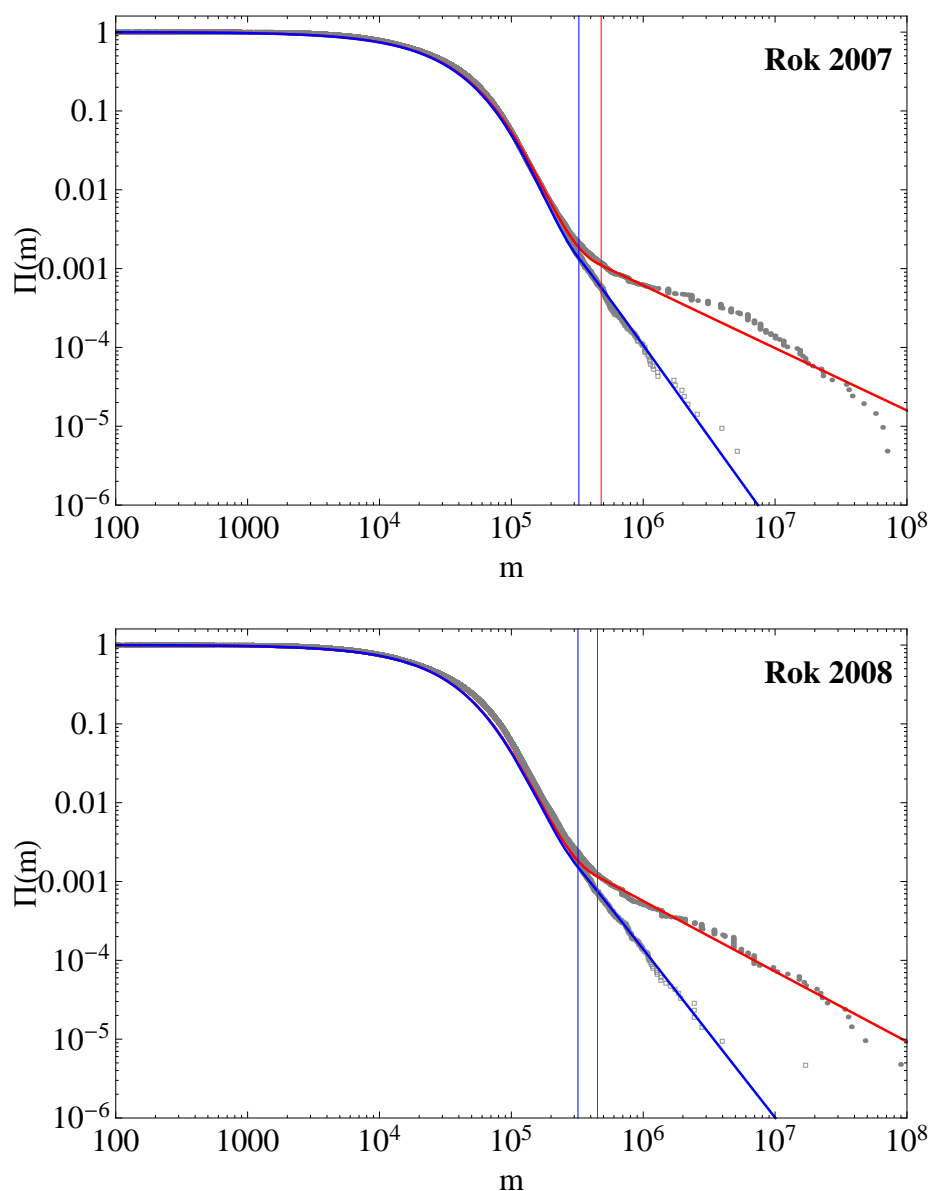
¹⁷ Zgodnie z rozważaniami w rozdziale 2.2 należy pamiętać, że $b = b'$ oraz $\alpha = \frac{a}{b} + 1$ i $\alpha_1 = \frac{a'}{b} + 1$.

Aby to pokazać, porównałem rozszerzony model Yakovenko z empirycznym dopełnieniem dystrybuanty (dotyczącym rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej) utworzonym tylko na podstawie danych Eurostatu (badanie EU-SILC)¹⁸. Otrzymane wyniki porównałem z dopasowaniami rozszerzonego modelu Yakovenko do empirycznych dopełnień dystrybuant zbudowanych przy użyciu połączonych baz danych – EU-SILC i magazynu Forbes. Wszystko to zamieściłem na rysunkach 4.14 – 4.16. Wartości parametrów pochodzące z obu porównań umieściłem w tabelach 4.8 – 4.10.

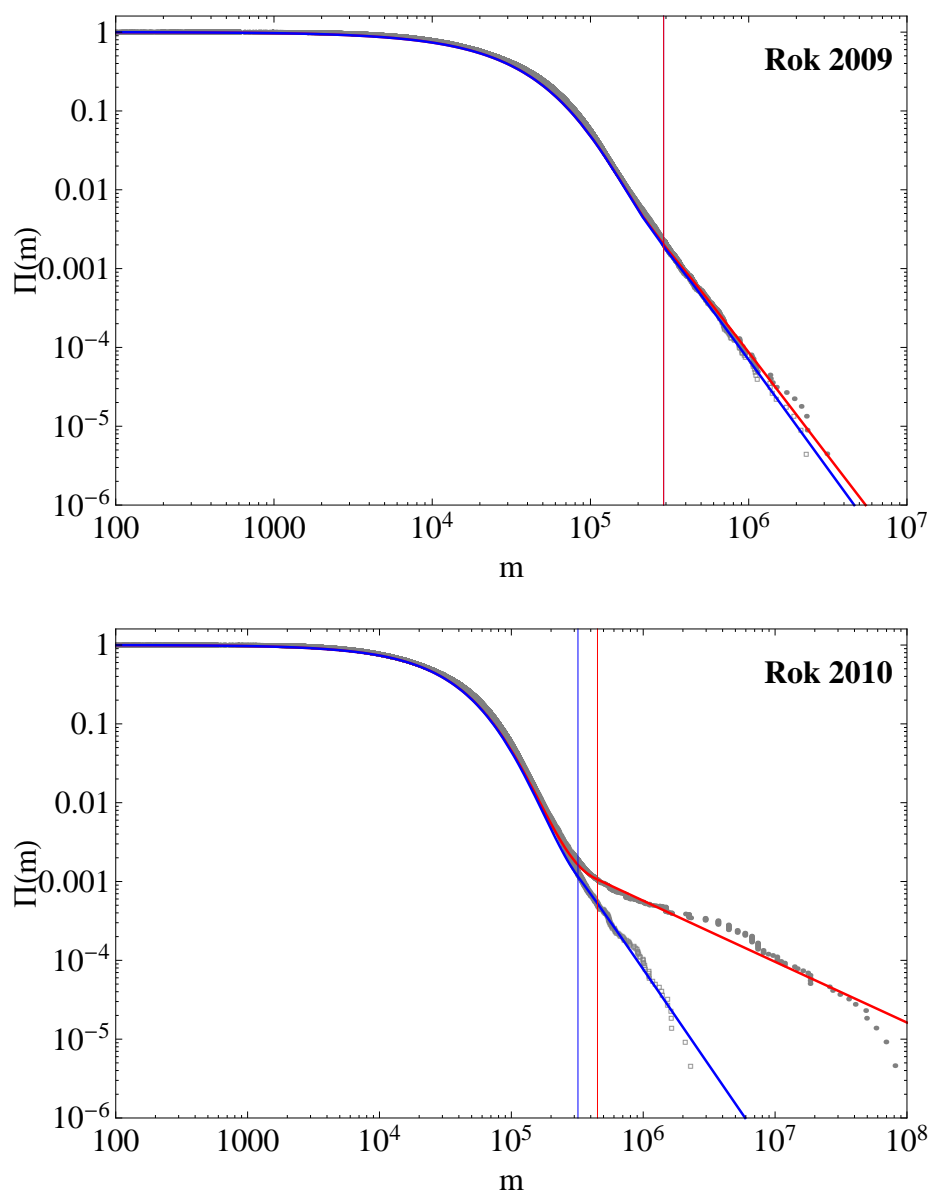
¹⁸ Szczegółowe wyniki analizy danych EU-SILC niezbędne do porównania rozszerzonego modelu Yakovenko z danymi empirycznymi Eurostatu zostały zamieszczone w dodatku F.



Rysunek 4.14: Porównanie, w skali log-log, rozszerzonego modelu Yakovenko (niebieska linia ciągła) z empirycznym dopełnieniem dystrybuanty rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej (baza danych EU-SILC; małe kwadraty) oraz porównanie rozszerzonego modelu Yakovenko (czerwona linia ciągła) z empirycznym dopełnieniem dystrybuanty rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej (połączona baza danych EU-SILC i magazynu Forbes; punkty) – lata 2005–2006. Niebieska i czerwona linia pionowa zaznaczają wartość m_1 , odpowiednio dla obu porównań.



Rysunek 4.15: Porównanie, w skali log–log, rozszerzonego modelu Yakovenko (niebieska linia ciągła) z empirycznym dopełnieniem dystrybuanty rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej (baza danych EU–SILC; małe kwadraty) oraz porównanie rozszerzonego modelu Yakovenko (czerwona linia ciągła) z empirycznym dopełnieniem dystrybuanty rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej (połączona baza danych EU–SILC i magazynu Forbes; punkty) – lata 2007–2008. Niebieska i czerwona linia pionowa zaznaczają wartość m_1 , odpowiednio dla obu porównań.



Rysunek 4.16: Porównanie, w skali log-log, rozszerzonego modelu Yakovenko (niebieska linia ciągła) z empirycznym dopełnieniem dystrybuanty rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej (baza danych EU-SILC; małe kwadraty) oraz porównanie rozszerzonego modelu Yakovenko (czerwona linia ciągła) z empirycznym dopełnieniem dystrybuanty rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej (połączona baza danych EU-SILC i magazynu Forbes; punkty) – lata 2009–2010. Niebieska i czerwona linia pionowa zaznaczają wartość m_1 , odpowiednio dla obu porównań. W przypadku roku 2009 linie te pokrywają się.

Tabela 4.8: Parametry T i T_1 uzyskane z porównania rozszerzonego modelu Yakovenko z empirycznym dopełnieniem dystrybuanty rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej – baza danych EU–SILC, oraz parametry T i T_1 uzyskane z porównania rozszerzonego modelu Yakovenko z empirycznym dopełnieniem dystrybuanty rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej – połączona baza danych EU–SILC i magazynu Forbes z lat 2005–2010.

Rok	Baza danych EU–SILC		Baza danych EU–SILC+Forbes	
	T	T_1	T	T_1
2005	$36\,000 \pm 3\,000$	$390\,000 \pm 50\,000$	$36\,000 \pm 3\,000$	$430\,000 \pm 50\,000$
2006	$37\,000 \pm 3\,000$	$330\,000 \pm 50\,000$	$37\,000 \pm 3\,000$	$445\,000 \pm 50\,000$
2007	$37\,000 \pm 3\,000$	$325\,000 \pm 50\,000$	$37\,000 \pm 3\,000$	$480\,000 \pm 50\,000$
2008	$38\,000 \pm 3\,000$	$320\,000 \pm 50\,000$	$38\,000 \pm 3\,000$	$450\,000 \pm 50\,000$
2009	$37\,000 \pm 3\,000$	$290\,000 \pm 50\,000$	$37\,000 \pm 3\,000$	$290\,000 \pm 50\,000$
2010	$38\,000 \pm 3\,000$	$320\,000 \pm 50\,000$	$38\,000 \pm 3\,000$	$450\,000 \pm 50\,000$

Tabela 4.9: Parametry m_0 i m_1 uzyskane z porównania rozszerzonego modelu Yakovenko z empirycznym dopełnieniem dystrybuanty rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej – baza danych EU–SILC, oraz parametry m_0 i m_1 uzyskane z porównania rozszerzonego modelu Yakovenko z empirycznym dopełnieniem dystrybuanty rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej – połączona baza danych EU–SILC i magazynu Forbes z lat 2005–2010.

Rok	Baza danych EU–SILC		Baza danych EU–SILC+Forbes	
	m_0 [EUR]	m_1 [EUR]	m_0 [EUR]	m_1 [EUR]
2005	$160\,000 \pm 20\,000$	$390\,000 \pm 50\,000$	$155\,000 \pm 20\,000$	$430\,000 \pm 50\,000$
2006	$150\,000 \pm 20\,000$	$330\,000 \pm 50\,000$	$145\,000 \pm 20\,000$	$445\,000 \pm 50\,000$
2007	$160\,000 \pm 20\,000$	$325\,000 \pm 50\,000$	$160\,000 \pm 20\,000$	$480\,000 \pm 50\,000$
2008	$120\,000 \pm 20\,000$	$320\,000 \pm 50\,000$	$120\,000 \pm 20\,000$	$450\,000 \pm 50\,000$
2009	$145\,000 \pm 20\,000$	$290\,000 \pm 50\,000$	$145\,000 \pm 20\,000$	$290\,000 \pm 50\,000$
2010	$140\,000 \pm 20\,000$	$320\,000 \pm 50\,000$	$135\,000 \pm 20\,000$	$450\,000 \pm 50\,000$

Tabela 4.10: Wykładniki α i α_1 uzyskane z porównania rozszerzonego modelu Yakovenko z empirycznym dopełnieniem dystrybuanty rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej – baza danych EU-SILC, oraz wykładniki α i α_1 uzyskane z porównania rozszerzonego modelu Yakovenko z empirycznym dopełnieniem dystrybuanty rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej – połączona baza danych EU-SILC i magazynu Forbes z lat 2005–2010.

Rok	Baza danych EU-SILC		Baza danych EU-SILC+Forbes	
	α	α_1	α	α_1
2005	3.216 ± 0.002	1.54 ± 0.02	2.907 ± 0.003	0.795 ± 0.009
2006	3.094 ± 0.003	2.15 ± 0.02	2.892 ± 0.004	0.86 ± 0.01
2007	3.057 ± 0.003	2.32 ± 0.01	2.735 ± 0.004	0.79 ± 0.01
2008	3.0632 ± 0.0005	2.13 ± 0.02	2.965 ± 0.001	0.890 ± 0.007
2009	2.979 ± 0.001	2.750 ± 0.005	2.974 ± 0.001	2.608 ± 0.006
2010	3.329 ± 0.001	2.43 ± 0.01	3.153 ± 0.002	0.77 ± 0.01

Analizując rysunki 4.14 – 4.16 oraz tabele 4.8 – 4.10 możemy zauważyć, że niezależnie od tego czy rozszerzony model Yakovenko porównujemy z danymi empirycznymi pochodzącymi z badania EU-SILC, czy z danymi empirycznymi powstałymi z połączenia baz EU-SILC i magazynu Forbes, parametry T , m_0 oraz α różnią się tylko nieznacznie. Zatem, kształt empirycznego dopełnienia dystrybuanty opisującego niskie i średnie dochody gospodarstw domowych w Unii Europejskiej w obu przypadkach jest bardzo zbliżony. Zmniejsza się natomiast wykładnik α_1 dla ogona rozkładu, opisującego wysokie dochody gospodarstw domowych przy jednoczesnym, nieznacznym rozszerzeniu zakresu empirycznego dopełnienia dystrybuanty odnoszącego się do średnich dochodów gospodarstw domowych (czyli wartości dochodów zawartych między m_0 i m_1).

Zatem, można powiedzieć, że połączenie bazy magazynu Forbes (zawierającej dane o wysokich dochodach gospodarstw domowych) z bazą danych EU-SILC (zawierającej dane o dochodach wszystkich trzech klas społecznych, tj. gospodarstw domowych uzyskujących niskie, średnie oraz wysokie dochody) wpływa tylko

i wyłącznie na ogony analizowanych empirycznych dopełnień dystrybuant, czyli na hierarchię wysokich dochodów gospodarstw domowych. Zastosowana procedura łączenia baz danych (patrz rozdział 4.2), zgodnie z oczekiwaniami, zwiększa liczbę obserwacji dotyczących gospodarstw domowych o wysokich dochodach, nie zmieniając struktury gospodarstw domowych o niskich i średnich dochodach¹⁹.

Należy podkreślić, że uwzględnienie w analizie danych empirycznych gospodarstw domowych o wysokich dochodach (poprzez odpowiednie połączenie danych EU-SILC i magazynu Forbes) ma istotny wpływ na otrzymane wyniki i czyni przedstawioną analizę znacznie pełniejszą.

Analiza porównawcza klas społecznych a współczynnik Giniego

Korzystając z rozszerzonego modelu Yakovenko prowadzącego do gęstości prawdopodobieństwa danej wzorem (4.3), można wyznaczyć współczynnik wyrażający względną liczebność klasy społecznej reprezentowanej przez gospodarstwa domowe o niskich dochodach w odniesieniu do klasy społecznej reprezentowanej przez gospodarstwa domowe o średnich dochodach. Współczynnik ten dany jest zależnością:

$$r_1 = \frac{\Pi(m_{\text{init}}) - \Pi(m_0)}{\Pi(m_0) - \Pi(m_1)}, \quad (4.4)$$

gdzie m_0 i m_1 są parametrami rozszerzonego modelu Yakovenko stanowiącymi graniczne wartości dochodu pomiędzy gospodarstwami domowymi, odpowiednio, o niskich i średnich dochodach oraz średnich i wysokich, natomiast m_{int} oznacza najmniejszą wartość dochodu gospodarstwa domowego, pojawiającą się w danych empirycznych.

Analogicznie można policzyć współczynnik wyrażający względną liczebność klasy społecznej reprezentowanej przez gospodarstwa domowe o średnich dochodach w odniesieniu do klasy społecznej reprezentowanej przez gospodarstwa domowe o wysokich dochodach. Zatem:

$$r_2 = \frac{\Pi(m_0) - \Pi(m_1)}{\Pi(m_1)}. \quad (4.5)$$

Wartości współczynników r_1 i r_2 zostały zamieszczone w tabeli 4.11.

¹⁹ Dodatkowe uwagi na temat procedury łączenia baz danych znajdują się w dodatku G.

Tabela 4.11: Współczynniki r_1 , r_2 oraz współczynnik Giniego (G) obliczone dla rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej – lata 2005–2010.

Rok	r_1	r_2	G [%]
2005	85.50	9.50	56.19
2006	63.25	13.01	54.34
2007	78.70	11.40	54.76
2008	37.79	22.33	54.94
2009	70.40	7.02	49.17
2010	54.91	16.86	55.42

Wartości współczynników r_1 i r_2 informują o względnej liczbie gospodarstw domowych należących do kolejnych klas społecznych. Wzrost współczynnika r_1 , przy jednoczesnym spadku współczynnika r_2 oznacza, że w społeczeństwie maleje liczba gospodarstw domowych o średnich dochodach. Część z tych gospodarstw (w stosunku do roku poprzedniego) zaczyna uzyskiwać niskie bądź wysokie dochody. Z kolei, spadek współczynnika r_1 , przy jednoczesnym wzroście współczynnika r_2 oznacza, że w społeczeństwie wzrasta liczba gospodarstw domowych o średnich dochodach. Innymi słowy, część z gospodarstw domowych o wysokich i niskich dochodach zaczyna uzyskiwać średnie dochody. Gdy ma miejsce wzrost obu współczynników wówczas wzrasta liczba gospodarstw domowych o niskich dochodach, maleje liczba gospodarstw domowych o wysokich dochodach a liczba gospodarstw o średnich dochodach pozostaje stała. Zatem, wzrost liczby gospodarstw domowych o niskich dochodach następuje kosztem klasy gospodarstw domowych o wysokich dochodach. Wreszcie, malenie obu współczynników oznacza spadek liczebności gospodarstw domowych o niskich dochodach i wzrost liczebności gospodarstw domowych o wysokich dochodach.

Analizując tabelę 4.11 możemy stwierdzić, że liczba gospodarstw domowych o średnich dochodach w kolejnych latach podlega naprzemiennym zmianom – raz wzrasta, a raz maleje. Przy czym, w roku 2009 nastąpił największy względny spadek liczby gospodarstw domowych uzyskujących średnie dochody.

Przedstawiona analiza cechuje się większą wrażliwością na zmiany w strukturze dochodów uzyskiwanych przez wszystkie trzy klasy społeczne, niż stosowany zwyczajowo współczynnik Giniego (G)²⁰ [113–117] (patrz tabela 4.11). Współczynnik ten różni się znacząco tylko dla roku 2009, kiedy to w wyniku kryzysu finansowego nastąpiła znacząca zmiana struktury rocznych dochodów (a więc i empirycznych dopełnień dystrybuant) uzyskiwanych przez gospodarstwa domowe w Unii Europejskiej. Niższa wartość współczynnika Giniego (tutaj $G=49.17$) oznacza, że w roku 2009 zmniejszyło się rozwarstwienie dochodów występujące w ramach społeczeństwa Unii Europejskiej. Zgadza się to dobrze z wcześniejszą obserwacją o degradacji w 2009 roku klasy gospodarstw domowych o wysokich dochodach.

Struktura procentowa dochodów gospodarstw domowych w Unii Europejskiej

Jako uzupełnienie rozważań dotyczących względnych liczebności klas społecznych r_1 i r_2 , wyznaczyłem podział procentowy rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej. Wyniki zamieściłem w tabeli 4.12.

Tabela 4.12: Struktura procentowa rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej – lata 2005–2010.

Rok	Względna liczba gospodarstw domowych o niskich dochodach $\Pi(m_{init}) - \Pi(m_0)$	Względna liczba gospodarstw domowych o średnich dochodach $\Pi(m_0) - \Pi(m_1)$	Względna liczba gospodarstw domowych o wysokich dochodach $\Pi(m_1)$
2005	98.72%	1.15%	0.13%
2006	98.33%	1.55%	0.12%
2007	98.64%	1.25%	0.11%
2008	97.31%	2.57%	0.12%
2009	98.40%	1.40%	0.20%
2010	98.11%	1.79%	0.10%

²⁰ Współczynnik Giniego został omówiony w dodatku H.

Na podstawie informacji zawartych w tabeli 4.12 możemy zauważyć, że gospodarstwa domowe o średnich i wysokich dochodach stanowią zaledwie ok. 1 – 3% wszystkich gospodarstw domowych w Unii Europejskiej. Jak widać, w roku 2008 nastąpił znaczący wzrost względnej liczby gospodarstw domowych o średnich dochodach. Jest to związane z rozszerzeniem się dla tego roku (patrz tabela 4.7) zakresu empirycznego dopełnienia dystrybucyjności odnoszącego się do średnich dochodów gospodarstw domowych (czyli wartości dochodów zawartych między m_0 i m_1). Z kolei, dla roku 2009 względna liczba gospodarstw domowych o wysokich dochodach jest największa. Przyczyną tego nie jest wzrost liczby tych gospodarstw domowych w roku 2009, a jedynie skokowy spadek wartości m_1 (spowodowany degradacją klasy gospodarstw domowych o wysokich dochodach; patrz tabela 4.7). Przez to znacząco poszerzeniu uległ zakres empirycznego dopełnienia dystrybucyjności odnoszącego się do wysokich dochodów gospodarstw domowych (czyli wartości dochodów większych niż m_1).

Znając całkowitą liczbę gospodarstw domowych w Unii Europejskiej, przy użyciu powyższej struktury procentowej (wyznaczonej za pomocą rozszerzonego modelu Yakovenko), można oszacować liczbę gospodarstw domowych uzyskujących odpowiednio niskie, średnie i wysokie dochody.

4.6 PODSUMOWANIE ROZDZIAŁU

W niniejszym rozdziale dokonano analizy empirycznych dopełnień dystrybucyjności dotyczących rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej w latach 2005–2010. Okazało się, że mogą one być dobrze opisywane dopełnieniem dystrybucyjności rozkładu wykładniczego (prawo Boltzmann–Gibbsa) w przypadku gospodarstw domowych o niskich dochodach, jak również słabym prawem Pareto dla gospodarstw uzyskujących średnie i wysokie dochody. **Wśród rozważanych modeli najbardziej obiecujące wyniki otrzymano dla rozszerzonego modelu Yakovenko. W modelu tym uzyskujemy teoretyczne dopełnienie dystrybucyjności opisujące dochody wszystkich trzech klas społecznych tj. gospodarstw domowych o niskich, średnich i wysokich dochodach, włączając w to obszary przejściowe pomiędzy tymi klasami.**

Tak szczegółowa analiza danych empirycznych była możliwa głównie dzięki proponowanej w pracy procedurze łączenia baz danych pochodzących z dwóch różnych źródeł, tj. badań EU-SILC i magazynu Forbes. Doprowadziła ona do uzyskania wystarczającej liczby danych empirycznych dla wszystkich trzech klas społecznych, tj. gospodarstw domowych uzyskujących niskie, średnie i wysokie dochody.

Wyznaczone parametry pochodzące z różnych modeli, a w szczególności z rozszerzonego modelu Yakovenko, mogą stanowić wskaźniki kryzysu. Kryzys nie wpływa na gospodarstwa domowe o niskich dochodach (parametry prawa Boltzmann–Gibbsa praktycznie nie zmieniają się), natomiast prowadzi do mniejszego rozwarstwienia społecznego wśród gospodarstw domowych o średnich i wysokich dochodach (oba wykładniki Pareto wyraźnie rosną). Warto jednak zauważyć, że w przypadku gospodarstw domowych o wysokich dochodach wykładnik Pareto doznaje wręcz skokowego wzrostu co oznacza, że skutki finansowe kryzysu dla najbogatszych obywateli społeczeństwa Unii Europejskiej były nadzwyczaj dotkliwe. Jednak w celu wysnucia dalszych wniosków (zwłaszcza dotyczących uniwersalnego charakteru uzyskanych wyników) niezbędne są dodatkowe badania, polegające przede wszystkim na przeprowadzeniu porównań z wcześniejszymi kryzysami.

Z kolei, klasa społeczna reprezentowana przez gospodarstwa domowe o niskich dochodach wydaje się zaskakująco stabilna. Jej dochody nie ulegają znaczącym zmianom w wyniku kryzysu. Zatem parametry rozkładów opisujących niskie dochody gospodarstw domowych nie mogą być indykatorami kryzysu. Jak się wydaje, tylko analiza średnich i wysokich dochodów gospodarstw domowych może dać odpowiedź na pytanie czy nadchodzi kryzys.

Reasumując, roczne dochody brutto gospodarstw domowych Unii Europejskiej mogą być opisane w całym zakresie za pomocą rozszerzonego modelu Yakovenko, obejmującego wszystkie klasy społeczne. Z kolei, dochody każdej klasy społecznej z osobna, mogą być również modelowane przy pomocy rozkładów:

- dopełnienia dystrybucyj rozkładu wykładniczego, w przypadku gospodarstw domowych o niskich dochodach (stanowiących około 98% całego społeczeństwa; patrz tabela 4.12),

- dopełnienia dystrybuanty rozkładu Pareto z wykładnikiem około 3, w przypadku gospodarstw domowych o średnich dochodach,
- dopełnienia dystrybuanty rozkładu Pareto z wykładnikiem około 1, w przypadku gospodarstw domowych o wysokich dochodach.

Należy podkreślić, że w analizie empirycznych dopełnień dystrybuant wykorzystałem dwa zasadnicze podejścia. Jedno dotyczy błędzeń losowych i pokazuje jaki mechanizm jest odpowiedzialny za pojawienie się prawa Pareto w przypadku gospodarstw domowych o wysokich dochodach. Drugie oparte na dynamice stochastycznej jest podejściem uniwersalnym w takim sensie, że przy jego pomocy opisałem dochody wszystkich klas społecznych.

ZAKOŃCZENIE

„A rich man is nothing but a poor man with money.”

WILLIAM CLAUDE FIELDS

Tematyka, którą podjąłem w niniejszej rozprawie doktorskiej stanowi aktywny obszar badań w dziedzinie ekonofizyki. Ich głównym celem jest pogłębione zrozumienie, a w tym opis ilościowy, mechanizmów uzyskiwania dochodów i bogacenia się lub ubożenia zarówno ludności Polski jak i Unii Europejskiej. Jest to niezwykle ważny cel, którego osiągnięcie pociąga za sobą możliwość określenia ścieżek rozwojowych Unii Europejskiej a w tym Polski.

Wkład niniejszej rozprawy do rozwoju wiedzy – modele dynamiki dochodów

W ramach przeprowadzonych rozważań i analiz wykorzystałem dwa zasadnicze podejścia. Jedno oparte jest na dynamice stochastycznej a drugie dotyczy błędzeń losowych.

Stosując pierwsze podejście udało mi się skonstruować model (nazwany przeze mnie rozszerzonym modelem Yakovenko) oparty na dynamice stochastycznej, opisujący dochody wszystkich trzech klas społecznych, tj. gospodarstw domowych o niskich, średnich i wysokich dochodach (patrz rozdział 2.2 oraz rozdział 4.4). Według mojej wiedzy jest to pierwszy model, będący w stanie opisać dynamikę zmian dochodu wszystkich trzech klas społecznych a więc całego społeczeństwa traktowanego jako jeden organizm. Dobra zgodność modelu z danymi empirycznymi została uzyskana dzięki przyjęciu następujących założeń:

- po pierwsze, rozszerzony model Yakovenko dopuszcza współistnienie procesu addytywnego i multiplikatywnego przy jednoczesnym zróżnicowaniu dynamiki zmian dochodu – w modelu tym przyjmuje się, że gospodarstwa domowe o niskich i średnich dochodach bogacą się (bądź tracą dochody) ina-

czej niż gospodarstwa domowe o wysokich dochodach (patrz rozdział 2.2, równanie (2.25) oraz (2.26)),

- po drugie, rozszerzony model Yakovenko spełnia warunek ciągłości gęstości prawdopodobieństwa (rozdział 2.2, równanie (2.27)).

W ramach drugiego podejścia zaproponowałem w pracy (patrz rozdział 2.1) model błędzenia losowego, który opisuje możliwy mechanizm bogacenia się gospodarstw domowych o wysokich dochodach prowadzący do słabego prawa Pareto, a więc prawa które dobrze opisuje dochody tej grupy gospodarstw domowych (patrz rozdział 3.3 oraz rozdział 4.4).

Wkład niniejszej rozprawy do rozwoju wiedzy – analiza danych empirycznych

W ramach niniejszej rozprawy doktorskiej dokonałem (według mojego rozeznania) najszerszej z dotychczas przeprowadzonych analiz danych empirycznych dotyczących rocznych dochodów rozporządzalnych gospodarstw domowych w Polsce oraz rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej. Analiza ta polegała w głównej mierze na weryfikacji omówionych w pracy modeli, tj. Prawa Efektów Proporcjonalnych (rozdział 1.6), prawa Boltzmanna–Gibbsa (rozdział 1.3), słabego prawa Pareto (rozdział 1.4), uogólnionego modelu Lotka–Volterra (rozdział 1.7) oraz, zaproponowanego przeze mnie, rozszerzonego modelu Yakovenko (rozdział 2.2). Dopiero rozszerzony model Yakovenko pozwala opisać dochody wszystkich klas społecznych; pozostałe z istniejących (wymienionych) modeli opisują jedynie dochody pojedynczych klas. W tabeli poniżej znajdują się informacje, przy pomocy jakich modeli mogą być opisywane dochody klas społecznych zarówno w Polsce jak i Unii Europejskiej (patrz rozdział 3.3 oraz rozdział 4.4).

Tabela : Modele opisujące dochody gospodarstw domowych w Polsce i Unii Europejskiej.

Gospodarstwa domowe	Polska	Unia Europejska
o niskich dochodach	Prawo Efektów Proporcjonalnych, uogólniony model Lotka–Volterra	Prawo Boltzmann–Gibbsa, rozszerzony model Yakovenko
o średnich dochodach	Słabe prawo Pareto, uogólniony model Lotka–Volterra	Słabe prawo Pareto, rozszerzony model Yakovenko
o wysokich dochodach	Słabe prawo Pareto	Słabe prawo Pareto, rozszerzony model Yakovenko

Należy podkreślić, że skonstruowanie i weryfikacja rozszerzonego modelu Yakovenko były możliwe, dzięki zaproponowaniu przeze mnie procedury łączenia baz danych. Procedura ta doprowadziła do uzyskania wystarczającej liczby danych o dochodach wszystkich trzech klas społecznych Unii Europejskiej (patrz rozdział 4.2 oraz dodatek G).

Istotnym wynikiem niniejszej rozprawy doktorskiej jest zauważenie, że wyznaczone parametry pochodzące z dopasowań rozważanych w pracy modeli teoretycznych do danych empirycznych, mogą stanowić wskaźniki kryzysu. Mianowicie, moje badania pokazują, że kryzys nie wpływa na gospodarstwa domowe o niskich dochodach zarówno w przypadku Polski jak i Unii Europejskiej (parametry Prawa Efektów Proporcjonalnych oraz prawa Boltzmann–Gibbsa praktycznie nie zmieniają się). Prowadzi on natomiast do mniejszego rozwarstwienia społecznego wśród gospodarstw domowych o średnich dochodach (wykładniki Pareto wyraźnie rosną). Na szczególną uwagę zasługuje fakt, że kryzys nie wpłynął negatywnie na polskie gospodarstwa domowe o wysokich dochodach (wykładnik Pareto ulegał niewielkim wahaniom). Był on jednak nadzwyczaj dotkliwy dla europejskich gospodarstw domowych o wysokich dochodach (skokowy wzrost wykładnika Pareto), powodując w praktyce zniknięcie w roku 2009 klasy społecznej reprezentowanej przez te gospodarstwa domowe.

Otwarte problemy badawcze

Choć przedstawiona przeze mnie analiza danych empirycznych ma charakter pogłębiony to jednak nadal pozostają otwarte następujące problemy badawcze:

- skonstruowanie modelu opisującego dochody wszystkich trzech klas społecznych w Polsce, tj. gospodarstw domowych o niskich, średnich i wysokich dochodach,
- opracowanie procedury łączenia baz danych (np. danych o dochodach pochodzących od Głównego Urzędu Statystycznego i danych o zamożności najbogatszych Polaków magazynu Wprost), dzięki której możliwe byłoby uzyskanie wystarczającej liczby danych o dochodach wszystkich trzech klas społecznych w Polsce,
- przeprowadzenie analizy porównawczej dochodów gospodarstw domowych Unii Europejskiej oraz Stanów Zjednoczonych.

Perspektywy wykorzystania wyników badań

Otrzymane rezultaty przyczynią się do lepszego zrozumienia mechanizmów bogacenia się gospodarstw domowych, klas społecznych oraz całych społeczeństw. Ponadto, otwiera się możliwość precyzyjnej klasyfikacji dochodowości gospodarstw domowych, tzn. określenia czy dane gospodarstwo domowe zalicza się do ubogiej, średniozamożnej bądź najbogatszej klasy społecznej. Wartości parametrów uzyskane z porównań danych empirycznych z modelami teoretycznymi mogą posłużyć do zdefiniowania bardziej zaawansowanych wskaźników charakteryzujących nierówności społeczne. Choć w naukach ekonomicznych do pomiaru nierówności społecznych stosuje się współczynnik Giniego, to jednak zaproponowana przeze mnie analiza oprócz potencjalnych wskaźników nierówności społecznych oferuje również cenne podejście teoretyczne na poziomie mikroskopowej dynamiki zmian dochodu poszczególnych gospodarstw domowych.

Uzyskane wyniki mogą posłużyć do skonstruowania efektywnej polityki społecznej mającej na celu redukcję nierówności społecznych głównie poprzez budowanie odpowiednich systemów podatkowych i emerytalnych. Będą one również

pomocne w analizie kształtowania się dochodów społeczeństw w czasie recesji, jak też w okresie rozwoju gospodarczego.

Podsumowując powyższą dyskusję wyników, mogę powiedzieć, że główne cele niniejszej rozprawy doktorskiej zostały osiągnięte.

WŁASNOŚCI ZMIENNYCH LOSOWYCH Z ROZKŁADU PARETO

Badaniem własności słabego rozkładu Pareto (wzór (1.61)) zajmował się m.in. Benoit Mandelbrot. Podał on kluczowe cechy zmiennych pochodzących z takiego rozkładu [6, 12].

Pierwsza, istotna własność mówi, iż rozkład Pareto (tutaj dla wykładnika $1 < \alpha < 2$) jest rozkładem stabilnym tzn., że zmienna losowa będąca sumą zmiennych losowych pochodzących z rozkładu Pareto sama podlega rozkładowi Pareto o tym samym wykładniku¹. Należy wspomnieć, że choć powyższą cechę posiadają również zmienne pochodzące z rozkładu Gaussa, to jednak dla rozkładu Pareto ma miejsce zasadnicza różnica – rozważane zmienne charakteryzują się nieskończoną wariancją (co nie ma miejsca w przypadku rozkładu Gaussa) [5, 6, 82].

Kolejna własność zmiennych losowych pochodzących ze słabego rozkładu Pareto związana jest z ideą rozkładów mieszanych (ang. *Weighted Mixture* [12]). Załóżmy, że nie wiemy nic o rozkładach, jakim podlegają dane o dochodach. Wówczas moglibyśmy przypuszczać, że dane te zostały uzyskane na „chybił trafił” z kilku różnych, bazowych rozkładów. W konsekwencji, rozkład opisujący obserwowane wartości dochodów byłby (ważoną) kombinacją rozkładów bazowych.

Rozważmy N niezależnych zmiennych losowych U_n podlegających rozkładowi opisanemu słabym prawem Pareto o wykładniku α :

$$\Pr(U_n > m) \sim C_n m^{-\alpha}, \quad \text{dla dużych } m, \quad (\text{A.1})$$

¹ Dowód został przedstawiony w dodatku B.

gdzie $\Pr(U_n > m)$ jest dopełnieniem dystrybuanty słabego prawa Pareto, a współczynnik C_n oznacza czynnik skalujący związany z wymiarem zmiennej (w ogólności) różny dla różnych zmiennych U_n , $n = 1, 2, \dots, N$.

Niech zmienna losowa U_W będzie średnią ważoną zmiennych U_n z wagami p_n ($U_W = \sum_{n=1}^N p_n U_n$, gdzie $\sum_{n=1}^N p_n = 1$). Wówczas rozkład zmiennej U_W będący (ważoną) kombinacją rozkładów Pareto przyjmie postać:

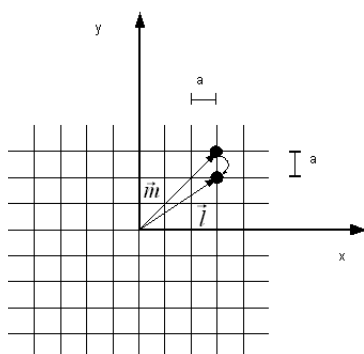
$$\Pr(U_W > m) = \Pr\left(\sum_{n=1}^N p_n U_n > m\right) = \sum_{n=1}^N p_n \Pr(U_n > m) \sim \sum_{n=1}^N C_n p_n m^{-\alpha} = C_W m^{-\alpha}. \quad (\text{A.2})$$

Tak więc U_W opisywane jest słabym prawem Pareto z czynnikiem skalującym $C_W = \sum_{n=1}^N C_n p_n$, będącym średnią ważoną czynników skalujących dla poszczególnych rozkładów Pareto [12].

Omówione powyżej przekształcenia Mandelbrota zmiennych losowych z rozkładu Pareto wskazują na cechy jakimi w ogólności powinien charakteryzować się rozkład niezmienniczy. Potwierdzają one, że prawo Pareto opisujące dochody ludności jest niewrażliwe na sposób, w jaki dane o tych dochodach są gromadzone i raportowane, co wynika z ich bezskalowego charakteru [5, 6].

ROZKŁAD SUMY N NIEZALEŻNYCH ZMIENNYCH LOSOWYCH POCHODZĄCYCH Z ROZKŁADU PARETO

Rozważmy błędzenie przypadkowe pojedynczej cząstki na sieci kwadratowej (rys. B.1). Przyjmujemy, że pojedyncze niezależne przemieszczenia (będące wektorowymi zmiennymi losowymi opisywanymi identycznym rozkładem prawdopodobieństwa) możliwe są tylko do najbliższych, sąsiednich węzłów sieci w otoczeniu cząstki.



Rysunek B.1: Błędzenie losowe cząstki na sieci dwuwymiarowej.

Uwzględniając powyższe założenia, możemy napisać, że po N krokach prawdopodobieństwo znalezienia cząstki w położeniu \vec{l} wynosi:

$$P_N(\vec{l}) = \sum_{\vec{m}} p(\vec{l} | \vec{m}) P_{N-1}(\vec{m}), \quad (\text{B.1})$$

gdzie $p(\vec{l} | \vec{m})$ jest prawdopodobieństwem pojedynczego przemieszczenia, a $P_{N-1}(\vec{m})$ wyraża prawdopodobieństwo znalezienia cząstki w położeniu \vec{m} po $N - 1$ kro-

kach. Zakładając (słabą) stacjonarność błędzenia przypadkowego $p(\vec{l} | \vec{m}) = p(\vec{l} - \vec{m})$ i zapisując równanie (B.1) w reprezentacji Fouriera otrzymujemy:

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} d\vec{k} e^{i\vec{k}\vec{l}} \tilde{P}_N(\vec{k}) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\vec{k} \sum_{\vec{m}} p(\vec{l} - \vec{m}) e^{i\vec{k}\vec{m}} \tilde{P}_{N-1}(\vec{k}). \quad (B.2)$$

Porównując czynniki pod całką możemy zapisać, że:

$$\tilde{P}_N(\vec{k}) = \sum_{\vec{m}} p(\vec{l} - \vec{m}) e^{-i\vec{k}(\vec{l} - \vec{m})} \tilde{P}_{N-1}(\vec{k}), \quad (B.3)$$

gdzie czynnik $\tilde{p}(\vec{k}) = \sum_{\vec{m}} p(\vec{l} - \vec{m}) e^{-i\vec{k}(\vec{l} - \vec{m})}$ nosi nazwę czynnika struktury błędzenia przypadkowego. Zatem, ostatecznie:

$$\tilde{P}_N(\vec{k}) = \tilde{p}(\vec{k}) \tilde{P}_{N-1}(\vec{k}). \quad (B.4)$$

Rozwiązując powyższe równanie rekurencyjnie otrzymujemy:

$$\tilde{P}_N(\vec{k}) = [\tilde{p}(\vec{k})]^N \tilde{P}_0(\vec{k}). \quad (B.5)$$

Ponieważ można przyjąć, że w chwili $t = 0$ cząstka będzie znajdowała się w ustalonym punkcie przestrzeni, to warunek początkowy możemy zapisać w postaci:

$$P_0(\vec{m}) = \delta(\vec{m}), \quad (B.6)$$

a stąd:

$$\tilde{P}_0(\vec{k}) = \int_{-\infty}^{\infty} d\vec{m} P_0(\vec{m}) e^{-i\vec{k}\vec{m}} = \int_{-\infty}^{\infty} d\vec{m} \delta(\vec{m}) e^{-i\vec{k}\vec{m}} = 1, \quad (B.7)$$

co po podstawieniu do (B.5) daje poszukiwaną ogólną formułę:

$$\tilde{P}_N(\vec{k}) = [\tilde{p}(\vec{k})]^N. \quad (B.8)$$

Zauważmy, że do wyprowadzenia powyższej formuły nie było konieczności podania jawnej postaci $\tilde{p}(\vec{k})$. Należy podkreślić, że wzór (B.8) pozostaje w mocy również w przypadku, gdy błędzenie cząstki ma charakter ciągły (w przestrzeni), tzn. możliwe są jej przemieszczenia do dowolnego punktu nieograniczonej przestrzeni euklidesowej.

Niech prawdopodobieństwo pojedynczego przemieszczenia cząstki będzie opisane rozkładem Lévy'ego postaci:

$$p(\vec{x}) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\vec{k}\vec{x} - \gamma|\vec{k}|^\alpha} d\vec{k}, \quad \alpha < 2, \quad (B.9)$$

gdzie \vec{x} oznacza przemieszczenie cząstki w przestrzeni euklidesowej. Można wykazać [6, 82], że dla $|\vec{x}| \rightarrow \infty$ gęstość rozkładu Pareto–Lévy’ego ma pogrubiony ogon, tzn.:

$$p(\vec{x}) \sim \frac{\alpha \gamma}{|\vec{x}|^{\alpha+1}}. \quad (\text{B.10})$$

Korzystając z formuły (B.8) wykazujemy poniżej stabilność rozkładu Lévy’ego, co jest równoznaczne z wykazaniem stabilności rozkładu Pareto.

Czynnik struktury błędzenia przypadkowego (czyli funkcja charakterystyczna rozkładu $p(x)$) wyraża się wzorem:

$$\tilde{p}(\vec{k}) = \int_{-\infty}^{\infty} p(\vec{x}) e^{-i\vec{k}\vec{x}} d\vec{x}. \quad (\text{B.11})$$

Dla rozkładu Lévy’ego $\tilde{p}(\vec{k})$ można znaleźć w prosty sposób zauważając, że w ogólności:

$$p(\vec{x}) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{p}(\vec{k}) e^{i\vec{k}\vec{x}} d\vec{k}, \quad (\text{B.12})$$

co po porównaniu ze wzorem (B.9) daje:

$$\tilde{p}(k) = e^{-\gamma|\vec{k}|^\alpha}. \quad (\text{B.13})$$

Otrzymaliśmy zatem funkcję charakterystyczną pojedynczej zmiennej losowej pochodzącej z rozkładu Lévy’ego. Z kolei dla sumy N niezależnych zmiennych losowych (N statystycznie niezależnych przemieszczeń cząstki) podlegających identycznemu rozkładowi Lévy’ego zgodnie ze wzorem (B.8) mamy:

$$\tilde{p}_N(\vec{k}) = \left[e^{-\gamma|\vec{k}|^\alpha} \right]^N = e^{-N\gamma|\vec{k}|^\alpha}, \quad (\text{B.14})$$

co w oczywisty sposób prowadzi do rozkładu Lévy’ego postaci:

$$P_N(\vec{x}) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\vec{k}\vec{x} - N\gamma|\vec{k}|^\alpha} d\vec{k}. \quad (\text{B.15})$$

oraz rozkładu Pareto:

$$P_N(\vec{x}) \sim \frac{\alpha N \gamma}{|\vec{x}|^{\alpha+1}}, \quad \text{przy } |\vec{x}| \rightarrow \infty. \quad (\text{B.16})$$

Wykazaliśmy zatem, że rozkład Pareto jest stabilny (ze względu na swój kształt) z czynnikiem normalizującym będącym sumą czynników normalizujących pojedynczych rozkładów [6, 82].

ELIMINACJA WSPÓŁCZYNNIKA $c(m_1, \dots, m_N, t)$
W UOGÓLNIONYM MODELU LOTKA-VOLTERRA

Dla infinitezymalnych przyrostów Δt (czyli $\Delta t \rightarrow dt$), równanie różnicowe (1.77) przechodzi w następujące równanie różniczkowe¹:

$$\frac{dm_i(t)}{dt} = [r_i(t) - 1]m_i(t) - c(m_1, \dots, m_N, t)m_i(t) + a m(t). \quad (C.1)$$

Sumując równanie (1.77) obustronnie po i oraz dzieląc obie strony przez N otrzymujemy:

$$\begin{aligned} \frac{1}{N} \sum_i m_i(t + \Delta t) - \frac{1}{N} \sum_i m_i(t) &= \frac{1}{N} \sum_i [r_i(t) - 1] \Delta t m_i(t) + \\ &- c(m_1, \dots, m_N, t) \Delta t \frac{1}{N} \sum_i m_i(t) + a \Delta t m(t). \end{aligned} \quad (C.2)$$

Z kolei po uwzględnieniu (1.80) mamy:

$$m(t + \Delta t) - m(t) = \frac{1}{N} \sum_i r_i(t) \Delta t m_i(t) - [c(m_1, \dots, m_N, t) + 1 - a] \Delta t m(t). \quad (C.3)$$

Postępując analogicznie, jak przy wyprowadzeniu równania (C.1) otrzymujemy:

$$\frac{dm(t)}{dt} = \frac{1}{N} \sum_i r_i(t) m_i(t) - [c(m_1, \dots, m_N, t) + 1 - a] m(t). \quad (C.4)$$

Wprowadzając nową zmienną:

$$x_i(t) = \frac{m_i(t)}{m(t)}, \quad \text{gdzie} \quad \sum_i x_i(t) = N, \quad (C.5)$$

¹ Aby otrzymać pochodną należało podzielić obie strony równania (1.77) przez Δt .

która wyraża względny dochód i -tego gospodarstwa domowego w chwili t , a następnie różniczkując ją:

$$\begin{aligned} \frac{dx_i(t)}{dt} &= \frac{1}{m(t)} \frac{dm_i(t)}{dt} - \frac{m_i(t)}{(m(t))^2} \frac{dm(t)}{dt} \\ \Leftrightarrow \frac{1}{m(t)} \frac{dm_i(t)}{dt} &= \frac{dx_i(t)}{dt} + \frac{m_i(t)}{(m(t))^2} \frac{dm(t)}{dt}, \end{aligned} \quad (C.6)$$

możemy równanie (C.1) przekształcić do postaci:

$$\begin{aligned} \frac{1}{m(t)} \frac{dm_i(t)}{dt} &= [r_i(t) - 1] \frac{m_i(t)}{m(t)} - c(m_1, \dots, m_N, t) \frac{m_i(t)}{m(t)} + a, \\ &= [r_i(t) - 1] x_i(t) - c(m_1, \dots, m_N, t) x_i(t) + a, \\ \frac{dx_i(t)}{dt} + \frac{m_i(t)}{(m(t))^2} \frac{dm(t)}{dt} &= [r_i(t) - 1] x_i(t) - c(m_1, \dots, m_N, t) x_i(t) + a, \\ \frac{dx_i(t)}{dt} + \frac{m_i(t)}{m(t)} \frac{1}{m(t)} \frac{dm(t)}{dt} &= [r_i(t) - 1] x_i(t) - c(m_1, \dots, m_N, t) x_i(t) + a, \\ \frac{dx_i(t)}{dt} + x_i(t) \frac{1}{m(t)} \frac{dm(t)}{dt} &= [r_i(t) - 1] x_i(t) - c(m_1, \dots, m_N, t) x_i(t) + a. \end{aligned} \quad (C.7)$$

Uwzględniając równanie (C.4) w postaci:

$$\begin{aligned} \frac{1}{m(t)} \frac{dm(t)}{dt} &= \frac{1}{N} \sum_i r_i(t) \frac{m_i(t)}{m(t)} - c(m_1, \dots, m_N, t) - 1 + a, \\ \frac{1}{m(t)} \frac{dm(t)}{dt} &= \frac{1}{N} \sum_i r_i(t) x_i(t) - c(m_1, \dots, m_N, t) - 1 + a \end{aligned} \quad (C.8)$$

i wstawiając do (C.7) dostajemy:

$$\frac{dx_i(t)}{dt} = r_i(t) x_i(t) + a - a x_i(t) - \frac{x_i(t)}{N} \sum_j r_j(t) x_j(t). \quad (C.9)$$

Otrzymaliśmy wyrażenie, które już nie zależy od $c(m_1, \dots, m_N, t)$. Tak więc mimo, iż pojedyncze gospodarstwa domowe są powiązane ze sobą za pomocą współczynnika $c(m_1, \dots, m_N, t)$, to ich względne dochody są już od niego niezależne.

Ostatni składnik w równaniu (C.9), łączący poszczególne indywidualne względne dochody, odgrywa kluczową rolę w wymianie bogactwa pomiędzy gospodarstwami domowymi. Ponieważ zmienne losowe $r_i(t)$ są od siebie statystycznie niezależne oraz nie zależą od zmiennych $x_i(t)$, to w granicy $N \rightarrow \infty$ możemy przyjąć, że:

$$\frac{1}{N} \sum_i r_i(t) x_i(t) \rightarrow 0. \quad (C.10)$$

A zatem ostatecznie otrzymujemy rozprzężony układ równań [6, 38–40]:

$$\frac{dx_i(t)}{dt} = r_i(t)x_i(t) + a - ax_i(t), \quad i = 1, \dots, N, \quad (\text{C.11})$$

stosunkowo łatwy do dalszej analizy.

Jeszcze raz należy podkreślić, że udało się zredukować uogólniony model Lotka–Volterra do postaci, w której nie występuje złożona zależność od współczynnika $c(m_1, \dots, m_N, t)$.

METODA WEIBULLA

Niech x_i ($i = 1, 2, \dots, n$) oznacza próbkę z rozkładu $P(x)$, o dopełnieniu dystrybuanty $\Pi(x)$, tworzącą szereg pozycyjny $x_1 \geq x_2 \geq \dots \geq x_n$. W przypadku bardzo dużej próbki, którą teoretycznie możemy uznać za nieskończoną, dopełnienie dystrybuanty $\Pi(x)$ określa ułamek danych w próbce mówiący, że wartość zmiennej losowej $x \geq x$. Wprowadźmy $n + 1$ przedziałów: $(-\infty; x_n]$, $(x_n; x_{n-1}]$, \dots , $(x_3; x_2]$, $(x_2; x_1]$, (x_1, ∞) . Wówczas na każdy przedział przypada ułamek $\frac{1}{n+1}$ całości próby. Zatem, ułamek $\frac{1}{n+1}$ jest estymatą $\Pi(x_1)$, $\frac{2}{n+1}$ jest estymatą $\Pi(x_2)$, itd., aż do ułamka $\frac{n}{n+1}$ będącego estymatą $\Pi(x_n)$. W ten sposób uzyskaliśmy empiryczne dopełnienie dystrybuanty $\tilde{\Pi}(x)$ (czyli estymatę dopełnienia dystrybuanty) w postaci funkcji schodkowej [95–97]:

$$\tilde{\Pi}(x) = \begin{cases} \frac{n}{n+1} & \text{dla } x \leq x_n, \\ \frac{i}{n+1} & \text{dla } x_{i+1} < x \leq x_i, \\ 0 & \text{dla } x > x_1. \end{cases} \quad (\text{D.1})$$

Wzór (D.1) jest wykorzystywany w pracy wszędzie tam, gdzie do analizy danych o dochodach gospodarstw domowych używa się empirycznego dopełnienia dystrybuanty.

OPIS ZMIENNEJ „ROCZNY DOCHÓD BRUTTO GOSPODARSTWA DOMOWEGO”

Zmienna – „roczny dochód brutto gospodarstwa domowego” analizowana w rozdziale 4 została skonstruowana w następujący sposób (cytuje w oryginale za [104]):

Total gross household income (HY010) is computed as:

The sum for all household members of gross personal income components:

- gross employee cash or near cash income (PY010G);
- company car (PY021G);
- gross cash benefits or losses from self-employment (including royalties) (PY050G);
- unemployment benefits (PY090G);
- old-age benefits (PY100G);
- survivor' benefits (PY110G);
- sickness benefits (PY120G);
- disability benefits (PY130G);
- education-related allowances (PY140G)

plus gross income components at household level:

- income from rental of a property or land (HY040G);
- family/children related allowances (HY050G);
- social exclusion not elsewhere classified (HY060G);

- housing allowances (HY070G);
- regular inter-household cash transfers received (HY080G);
- interests, dividends, profit from capital investments in unincorporated business (HY090G);
- income received by people aged under 16 (HY110G).

That means:

$$HY010 = HY040G + HY050G + HY060G + HY070G + HY080G + HY090G + HY110G + [\text{for all household members}](PY010G + PY021G + PY050G + PY090G + PY100G + PY110G + PY120G + PY130G + PY140G)^1.$$

¹ Szczegółowy opis występujących tutaj zmiennych znajduje się w dokumentacji Eurostatu [104].

ANALIZA DANYCH EMPIRYCZNYCH – BADANIE EU-SILC

Tabela F.1: Parametry uzyskane z dopasowania dopełnienia dystrybuanty rozkładu wykładniczego (prawo Boltzmann–Gibbsa) do empirycznego dopełnienia dystrybuanty rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej (dane EU-SILC) z lat 2005–2010.

Rok	T
2005	$35\,597 \pm 5$
2006	$35\,288 \pm 4$
2007	$36\,802 \pm 5$
2008	$34\,873 \pm 3$
2009	$36\,163 \pm 5$
2010	$35\,168 \pm 4$

Tabela F.2: Parametry uzyskane z dopasowania słabego prawa Pareto do środkowej części empirycznego dopełnienia dystrybuanty rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej (dane EU–SILC) z lat 2005–2010.

Rok	α	m_0 [EUR]
2005	3.216 ± 0.002	$39\,767 \pm 47$
2006	3.094 ± 0.003	$37\,546 \pm 54$
2007	3.057 ± 0.003	$37\,801 \pm 53$
2008	3.0632 ± 0.0005	$39\,497 \pm 8$
2009	2.979 ± 0.001	$36\,901 \pm 26$
2010	3.329 ± 0.001	$42\,367 \pm 18$

Tabela F.3: Parametry uzyskane z analizy rankingów rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej (dane EU–SILC) z lat 2005–2010.

Rok	α_{ranking1}	Const	$\alpha_{\text{Pareto1}} = 1/\alpha_{\text{ranking1}}$
2005	0.648 ± 0.009	15.88 ± 0.03	1.54 ± 0.02
2006	0.465 ± 0.004	15.23 ± 0.02	2.15 ± 0.02
2007	0.430 ± 0.002	15.14 ± 0.01	2.32 ± 0.01
2008	0.470 ± 0.004	15.42 ± 0.02	2.13 ± 0.02
2009	0.3636 ± 0.0007	14.829 ± 0.004	2.750 ± 0.005
2010	0.411 ± 0.002	14.98 ± 0.01	2.43 ± 0.01

DODATKOWE UWAGI NA TEMAT PROCEDURY ŁĄCZENIA BAZ DANYCH

W rozdziale 4.2 przedstawiono procedurę łączenia baz danych, która doprowadziła do uzyskania wystarczającej liczby danych empirycznych dla wszystkich trzech klas społecznych Unii Europejskiej, tj. gospodarstw domowych uzyskujących niskie, średnie i wysokie dochody. Procedura ta polegała na odpowiednim zmodyfikowaniu danych dotyczących wartości majątków najbogatszych Europejczyków (magazyn Forbes) i połączeniu ich z danymi o dochodach pochodzącymi z bazy EU-SILC. Chociaż dane magazynu Forbes stanowią przybliżone oszacowanie majątków posiadanych przez najbogatszych, to pozwalają one dość dobrze ustalić ich hierarchię wyrażoną poprzez ranking dochodów.

Procedura łączenia baz danych została zaproponowana głównie dlatego, że użyte do analizy bazy danych, choć jedyne dostępne, są niepełne. Niestety baza danych EU-SILC zawiera bardzo mało obserwacji dotyczących gospodarstw domowych o wysokich dochodach. Aby temu zaradzić wykorzystano w niniejszej rozprawie dane magazynu Forbes. Należy podkreślić, że celem skonstruowanej procedury było, oprócz uzyskania bazy danych zawierającej dane empiryczne w postaci dochodów wszystkich klas społecznych, odtworzenie „ogonów” dopełnień dystrybuant (dotyczących gospodarstw domowych o wysokich dochodach, scharakteryzowanych przez wykładnik Pareto α_1) analizowanych w rozdziale 4. W rozdziale 4.4 wykładniki Pareto zostały wyznaczone przy użyciu połączonej bazy danych EU-SILC i magazynu Forbes (patrz sekcja pt.: *Rankingi gospodarstw domowych o średnich i wysokich dochodach*). Z kolei, w rozdziale 4.5 wyznaczono ich wartości wyłącznie dla bazy danych EU-SILC (patrz sekcja pt.: *Rozszerzony model Yakovenko a procedura*

łączenia baz danych). Zauważono, że wartości tych wykładników są wówczas dużo wyższe niż w przypadku połączonych baz (dla których wykładnik Pareto był bliski jedności). Skąd zatem wiadomo, że należy łączyć bazy tak, aby uzyskać niższą wartość wykładnika Pareto? Wynika to z faktu, jak pokazano w rozdziałach 3.3 i 4.4, że wykładnik Pareto opisujący ogony dopełnień dystrybuant można oszacować przy pomocy analizy rankingu. Jeśli utworzymy ranking najbogatszych Europejczyków wykorzystując tylko i wyłącznie dane magazynu Forbes to okaże się, że wykładniki Pareto są bliskie jedności. Zatem ogony analizowanych dopełnień dystrybuant również powinny być bliskie jedności, gdyż jak pamiętamy $\alpha_{\text{pareto1}} = \frac{1}{\alpha_{\text{ranking1}}}$. W tabeli G.1 zamieszczono wykładniki Pareto α_1 wyznaczone przy użyciu połączonych baz danych EU-SILC i magazynu Forbes, oraz wyznaczone przy użyciu wyłącznie bazy danych magazynu Forbes.

Tabela G.1: Wykładniki Pareto α_1 wyznaczone przy użyciu połączonych baz danych EU-SILC i magazynu Forbes oraz wykładniki Pareto α_1 wyznaczone tylko przy użyciu bazy danych magazynu Forbes – lata 2005–2010. W przypadku danych magazynu Forbes oszacowanie wykładnika Pareto dla roku 2009 nie było możliwe, gdyż w bazie mamy tylko 3 obserwacje (patrz rozdział 4).

	Baza danych EU-SILC+Forbes	Baza danych Forbes
Rok	α_1	α_1
2005	0.795 ± 0.009	0.79 ± 0.02
2006	0.86 ± 0.01	0.94 ± 0.04
2007	0.79 ± 0.01	0.90 ± 0.03
2008	0.890 ± 0.007	0.83 ± 0.03
2009	2.608 ± 0.006	–
2010	0.77 ± 0.01	0.86 ± 0.03

Jak widać, wartości wyznaczonych niezależnie wykładników Pareto są zbliżone. Możliwe przyczyny pojawiających się niewielkich różnic zostaną omówione w dalszej części rozważań.

Jedno z pytań jakie pojawia się przy łączeniu baz danych, to pytanie o wpływ łączenia baz danych na kształt analizowanych dopełnień dystrybuant. Czy kształt otrzymanych dopełnień dystrybuant po połączeniu baz danych jest taki sam jak kształt dopełnień, które otrzymalibyśmy posiadając hipotetyczną, pełną bazę zawierającą dane o dochodach wszystkich gospodarstw domowych w Unii Europejskiej? Aby odpowiedzieć na to pytanie przeprowadzono symulację komputerową. W tym celu wylosowano dwie niezależne próby liczb z rozkładu Pareto opisywanego gęstością prawdopodobieństwa postaci¹:

$$P(m) = \frac{\alpha \lambda^\alpha}{(\lambda + m)^{\alpha+1}}, \quad m \geq 0, \alpha, \lambda > 0, \quad (G.1)$$

dla którego dopełnienie dystrybuanty wyraża się wzorem:

$$\Pi(m) = \left(\frac{\lambda}{\lambda + m} \right)^\alpha, \quad m \geq 0, \alpha, \lambda > 0. \quad (G.2)$$

Pierwsza próba, N , zawierała 5×10^6 liczb z rozkładu Pareto o wykładniku $\alpha = 3$; druga próba, N_1 zawierała także 5×10^6 liczb z rozkładu Pareto o wykładniku $\alpha_1 = 1$. Wykładniki te dobrano tak, aby odtworzyć nachylenia występujące dla dopełnień dystrybuant utworzonych na podstawie danych empirycznych. Z kolei, parametr $\lambda = 1$ dla pierwszej próby oraz $\lambda_1 = 1000$ dla drugiej, co pozwala traktować wylosowane liczby jako hipotetyczny dochód gospodarstw domowych.

W kolejnym kroku z wylosowanej próby N_1 gospodarstw domowych wybrano próbkę S zawierającą 200 najbogatszych gospodarstw domowych. Otrzymana w ten sposób próbka imituje sytuację, w której posiadamy dane o dochodach najbogatszych Europejczyków uzyskanych z magazynu Forbes. W próbie S możemy wyróżnić najniższy „dochód”, który oznaczamy jako m_{\min} .

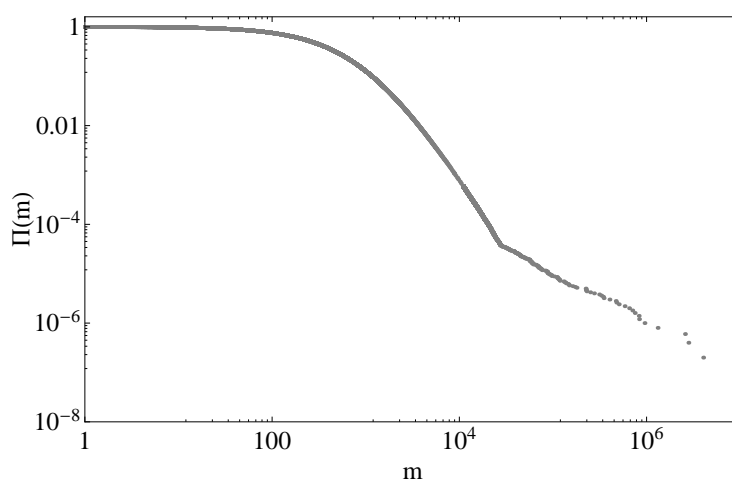
Następnie, z próby N usunięto gospodarstwa domowe², których dochód był większy niż m_{\min} (tę nową próbę oznaczmy przez N_3), po czym połączono ją z próbą S . W ten sposób odtworzono sytuację pojawiającą się dla danych empirycznych, gdzie gospodarstwa domowe o średnich dochodach opisywane są sła-

¹ Rozkład ten wybrano dlatego, aby z jednej strony odtworzyć sytuację podobną do tej jaką obserwujemy dla danych empirycznych dotyczących dochodów gospodarstw domowych w Unii Europejskiej a z drugiej otrzymać rozkłady potęgowe umożliwiające, w dalszym toku rozważań kontrolę nachylenia ogonów dopełnień dystrybuant.

² Liczba usuniętych gospodarstw domowych wyniosła 225.

bym prawem Pareto z wykładnikiem $\alpha = 3$ a gospodarstwa domowe o wysokich dochodach opisywane są słabym prawem Pareto z wykładnikiem $\alpha_1 = 1$.

Utworzoną w ten sposób nową próbę $N_3 + S$ traktujemy jako dane o dochodach wszystkich gospodarstw domowych Unii Europejskiej. Dopełnienie dystrybuanty, utworzone (przy pomocy metody Weibulla – patrz rozdział 3.2) na podstawie tej nowej próby znajduje się na rysunku G.1 (skala log–log). Niestety, wykonanie analogicznego wykresu dla danych empirycznych nie jest możliwe, gdyż nie dysponujemy bazą zawierającą dane o dochodach wszystkich gospodarstw domowych w Unii Europejskiej.

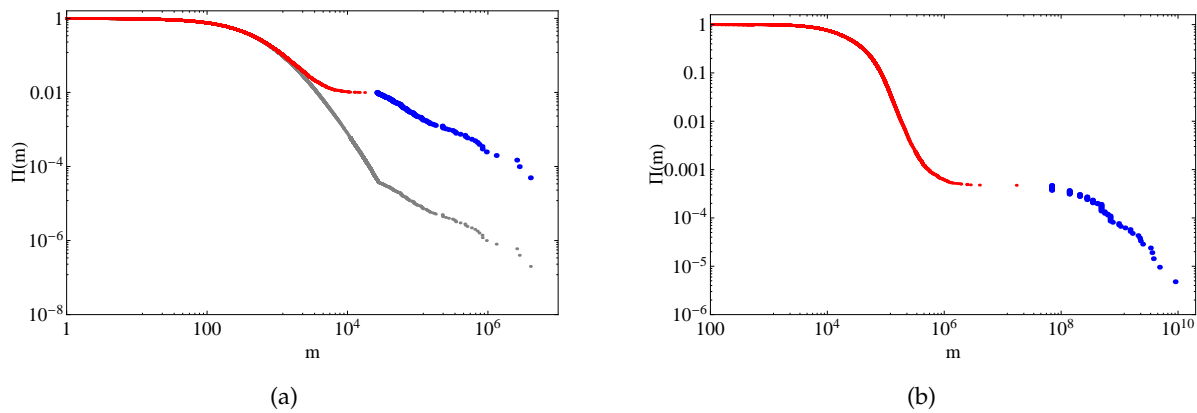


Rysunek G.1: Dopełnienie dystrybuanty utworzone na podstawie próby $N_3 + S$ liczb wylosowanych z rozkładów Pareto opisywanych równaniem (G.1). Wylosowane liczby reprezentują dochody gospodarstw domowych w Unii Europejskiej.

Następnie z próby N_3 wybrano losowo próbę P o liczebności 20 000 gospodarstw domowych. W tym przypadku otrzymana próba imituje sytuację, w której posiadamy dane ankietowe pochodzące z bazy EU–SILC.

Dalej postępujemy tak samo jak w przypadku danych empirycznych Eurostatu i magazynu Forbes (patrz rozdział 4.2, trzeci punkt procedury łączenia baz danych). Mianowicie, łączymy próbę S z próbą P . Następnie, dla tak utworzonej próby konstruujemy dopełnienie dystrybuanty (rysunek G.2a, skala log–log.). Dla porównania na rysunku G.2b znajduje się empiryczne dopełnienie dystrybuanty rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej z roku 2008, przed zastosowaniem ostatniego punktu procedury łączenia baz danych (patrz rozdział

4.2), czyli przed pomnożeniem danych o dochodach z magazynu Forbes przez czynnik proporcjonalności (wynoszący 1.0×10^{-2}).

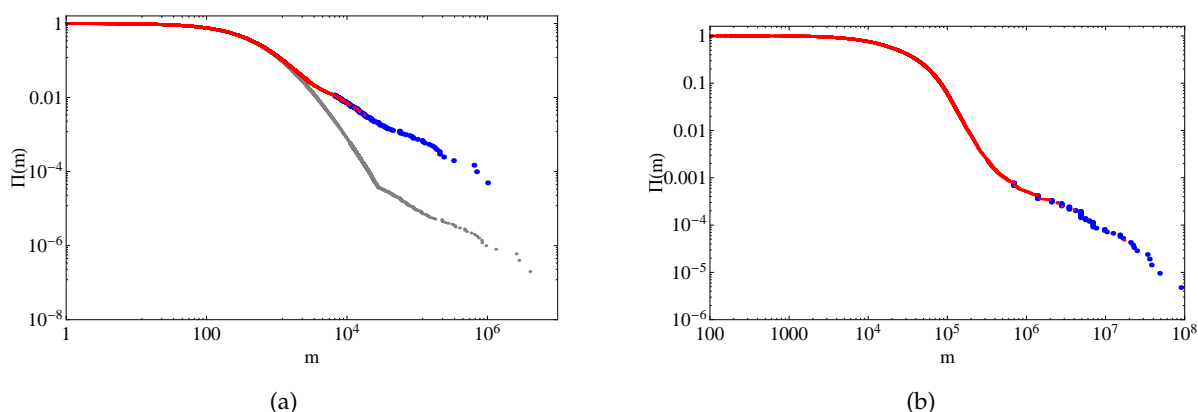


Rysunek G.2: (a) Dopełnienie dystrybuanty (czerwone i niebieskie punkty) utworzone na podstawie połączonej próby $P + S$ liczb wylosowanych z rozkładów Pareto opisanych równaniem (G.1). Wylosowane liczby reprezentują dochody gospodarstw domowych w Unii Europejskiej. Jak widać, otrzymany rozkład jest nieciągły. Czerwone punkty odpowiadają danym z próby P a niebieskie danym z próby S . Na rysunku zamieszczono dla porównania dopełnienie dystrybuanty (szare punkty) utworzone na podstawie próby $N_3 + S$. (b) Empiryczne dopełnienie dystrybuanty rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej z roku 2008. Punkty czerwone reprezentują dane empiryczne pochodzące z badania EU-SILC a niebieskie dane empiryczne pochodzące z magazynu Forbes.

Jak możemy zauważyć, otrzymane rozkłady są nieciągłe. Nieciągłość ta wynika z faktu, że zarówno w przypadku danych z symulacji jak i danych empirycznych mamy pewne niedoszacowanie dochodów gospodarstw domowych zawartych pomiędzy najbogatszymi z badania EU-SILC (lub odpowiednio próby P) i „najbiedniejszymi” z rankingu magazynu Forbes (lub odpowiednio z próby S). Należy też zwrócić uwagę na fakt, że choć otrzymane dopełnienie dystrybuanty (rysunek G.2a) ma inny kształt niż dopełnienie przedstawione na rysunku G.1, to w obu przypadkach nachylenie ogonów jest identyczne.

Aby pozbyć się nieciągłości występującej dla analizowanych dopełnień dystrybuant, postąpiono analogicznie jak w czwartym punkcie procedury łączenia baz danych (patrz rozdział 4.2). Dane o dochodach z próby S pomnożono przez czynnik proporcjonalności, który wyniósł 0.25 (przypomnijmy że dla danych empirycznych

nych dotyczących dochodów gospodarstw domowych w Unii Europejskiej czynnik ten wyniósł 1.0×10^{-2}) a następnie próbę tę połączono z próbą P. Wykres dopełnienia dystrybuanty otrzymanego dla zmodyfikowanej próby P + S znajduje się na rysunku G.3a (skala log-log). Dla porównania na rysunku G.3b znajduje się empiryczne dopełnienie dystrybuanty rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej z roku 2008, otrzymane po pomnożeniu danych z magazynu Forbes przez czynnik proporcjonalności.



Rysunek G.3: (a) Dopełnienie dystrybuanty (czerwone i niebieskie punkty) utworzone na podstawie zmodyfikowanej próby P + S. Tym razem otrzymany rozkład jest ciągły. Czerwone punkty odpowiadają danym z próby P a niebieskie danym z próby S pomnożonej przez czynnik proporcjonalności. Na rysunku zamieszczono dla porównania dopełnienie dystrybuanty (szare punkty) utworzone na podstawie próby $N_3 + S$. (b) Empiryczne dopełnienie dystrybuanty rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej z roku 2008. Punkty czerwone reprezentują dane empiryczne pochodzące z badania EU-SILC a niebieskie dane empiryczne pochodzące z magazynu Forbes (po pomnożeniu przez czynnik proporcjonalności).

Analizując wykres G.3a możemy zauważyć, że zastosowanie czynnika skalującego do danych z próby S uczyniło analizowane dopełnienie dystrybuanty rozkładem ciągłym, jednocześnie zachowując nachylenie ogona, który nadal opisywany jest prawem Pareto z wykładnikiem $\alpha_1 = 1$. Dopełnienie to ma kształt podobny do dopełnienia dystrybuanty utworzonego na podstawie próby $N_3 + S$ (patrz rysunek G.1). Zniekształceniu uległa wyłącznie środkowa część rozkładu. Wynika to przede wszystkim z faktu, że dla próby P + S największe niedoszacowanie występuje dla gospodarstw domowych uzyskujących średnie dochody.

Podsumowując: zastosowanie czynnika skalującego w procedurze łączenia baz danych pozwala z dobrym przybliżeniem odtworzyć dopełnienia dystrybuant jakie otrzymalibyśmy posiadając dane o dochodach wszystkich gospodarstw domowych w Unii Europejskiej. Czynniki ten dobieramy tak, aby z jednej strony, po połączeniu baz danych, otrzymać rozkłady ciągłe a z drugiej strony odtworzyć poprawne nachylenie ich ogonów. Dobór zbyt małego czynnika skalującego doprowadziłby do niepoprawnych nachyleń ogonów dopełnień dystrybuant. Z kolei, czynnik zbyt duży nadal dawałby rozkłady nieciągłe.

W przypadku danych empirycznych dotyczących rocznych dochodów brutto gospodarstw domowych w Unii Europejskiej wartość czynnika skalującego (na poziomie 1.0×10^{-2}) pozwala również na przybliżone oszacowanie jaką część różnicy majątków stanowi dochód.

WSPÓŁCZYNNIK GINIEGO

Współczynnik Giniego (nazywany również indeksem Giniego lub wskaźnikiem Giniego) jest statystyczną miarą dyspersji zaproponowaną przez włoskiego statystyka i socjologa Corrado Giniego. Wyraża on nierównomierność rozkładu zmiennej losowej [117].

Współczynnik Giniego jest powszechnie stosowany jako miara nierówności uzyskiwanych przez gospodarstwa domowe dochodów lub posiadanego przez nie bogactwa. Jeśli wartości dochodów m_i uzyskiwanych przez poszczególne gospodarstwa domowe uporządkujemy rosnąco ($m_i < m_{i+1}$) to wówczas współczynnik Giniego wyraża się wzorem:

$$G = \frac{\sum_{i=1}^N (2i - N - 1) m_i}{N^2 \langle m \rangle}, \quad (\text{H.1})$$

gdzie N jest liczbą gospodarstw domowych w analizowanej próbie a $\langle m \rangle = \frac{\sum_{i=1}^N m_i}{N}$.

Współczynnik Giniego zawiera się w przedziale od zera (gdy wszystkie gospodarstwa domowe uzyskują taki sam dochód) do jeden (gdy tylko jedno gospodarstwo domowe uzyskuje dochody, a pozostałe nie osiągają żadnych dochodów) [113–117].

BIBLIOGRAFIA

- [1] H.E. Stanley, V. Afanasyev, L.A.N. Amaral, S.V. Buldyrev, A.L. Goldberger, S. Havlin, H. Leschhorn, P. Maass, R.N. Mantegna, C.-K. Peng, P.A. Prince, M.A. Salinger, M.H.R. Stanley, G.M. Viswanathan, *Anomalous fluctuations in the dynamics of complex systems: from DNA and physiology to econophysics*, Physica A 224 (1996), str. 302–321.
- [2] V.M. Yakovenko, J.B. Rosser, *Colloquium: Statistical mechanics of money, wealth, and income*, Reviews of Modern Physics 81 (2009), str. 1703–1725.
- [3] V. Pareto, *Cours d'économie politique*, L'Université de Lausanne, 1897.
- [4] B. Mandelbrot, *The Pareto–Lévy Law and the Distribution of Income*, International Economic Review 1 (1960), str. 79–106.
- [5] P. Richmond, S. Hutzler, R. Coelho, P. Repetowicz, *A review of empirical studies and models of income distributions in society*, w książce: *Econophysics & Sociophysics: Trends & Perspectives*, pod redakcją: B.K. Chakrabarti, A. Chakraborti, A. Chatterjee, WILEY-VCH, Weinheim, 2006, str. 129–158.
- [6] M. Jagielski, *Badanie zamożności gospodarstw domowych w Polsce metodami egzotycznej i tradycyjnej fizyki statystycznej*, praca magisterska, Wydział Fizyki, Uniwersytet Warszawski, Warszawa, 2009, dostępna w bibliotece Instytutu Fizyki Doświadczalnej.
- [7] R. Gibrat, *Les inégalités économiques: applications aux inégalités des richesses, à la concentration des entreprises...d'une loi nouvelle, la loi de l'effet proportionnel*, Recueil Sirey, Paris, 1931.
- [8] M. Kalecki, *On the Gibrat Distribution*, Econometrica 13 (1945), str. 161–170.
- [9] M. Armatte, *Robert Gibrat et la loi de l'effet proportionnel*, Mathématiques et sciences humaines 129 (1995), str. 5–35.

- [10] J. Sutton, *Gibrat's Legacy*, Journal of Economic Literature 35 (1997), str. 40–59.
- [11] D.G. Champernowne, *A Model of Income Distribution*, The Economic Journal 63 (1953), str. 318–351.
- [12] B. Mandelbrot, *New Methods in Statistical Economics*, The Journal of Political Economy 71 (1963), str. 421–440.
- [13] J. Angle, *The Surplus Theory of Social Stratification and the Size Distribution of Personal Wealth*, Social Forces 65 (1986), str. 293–326.
- [14] J. Angle, *The Inequality Process and the Distribution of Income to Blacks and Whites*, Journal of Mathematical Sociology 17 (1992), str. 77–98.
- [15] J. Angle, *Deriving the Size Distribution of Personal Wealth from "The Rich Get Richer, the Poor Get Poorer"*, Journal of Mathematical Sociology 18 (1993), str. 27–46.
- [16] J. Angle, *How the Gamma Law of Income Distribution Appears Invariant Under Aggregation*, Journal of Mathematical Sociology 21 (1996), str. 325–358.
- [17] S. Ispolatov, P.L. Krapivsky, S. Redner, *Wealth distributions in asset exchange models*, The European Physical Journal B 2 (1998), str. 267–276.
- [18] A. Chakraborti, B.K. Chakrabarti, *Statistical mechanics of money: how saving propensity affects its distribution*, The European Physical Journal B 17 (2000), str. 167–170.
- [19] A. Drăgulescu, V.M. Yakovenko, *Statistical mechanics of money*, The European Physical Journal B 17 (2000), str. 723–729.
- [20] J. Angle, *The Statistical Signature of Pervasive Competition on Wage and Salary Incomes*, Journal of Mathematical Sociology 26 (2002), str. 217–270.
- [21] J. Angle, *The Inequality Process as a wealth maximizing process*, Physica A 367 (2006), str. 388–414.
- [22] A. Chatterjee, B.K. Chakrabarti, *Kinetic exchange models for income and wealth distributions*, The European Physical Journal B 60 (2007), str. 135–149.

- [23] R. Fischer, D. Braun, *Transfer potentials shape and equilibrate monetary systems*, Physica A 321 (2003), str. 605–618.
- [24] N. Xi, N. Ding, Y. Wang, *How required reserve ratio affects distribution and velocity of money*, Physica A 357 (2005), str. 543–555.
- [25] P. Cockshott, A. Cottrell, *Probabilistic political economy and endogenous money*, 2008, praca niepublikowana, dostępna pod adresem: <http://www.dcs.gla.ac.uk/publications/PAPERS/8935/probpolecon.pdf>.
- [26] M. Patriarca, A. Chakraborti, K. Kaski, *Statistical model with a standard Γ distribution*, Physical Review E 70 (2004), str. 016104.
- [27] M. Patriarca, A. Chakraborti, K. Kaski, *Gibbs versus non-Gibbs distributions in money dynamics*, Physica A 340 (2004), str. 334–339.
- [28] A. Chatterjee, B.K. Chakrabarti, S.S. Manna, *Pareto law in a kinetic model of market with random saving propensity*, Physica A 335 (2004), str. 155–163.
- [29] J. C. Ferrero, *The statistical distribution of money and the rate of money transference*, Physica A 341 (2004), str. 575–585.
- [30] N. Scafetta, S. Picozzi, B.J. West, *An out-of-equilibrium model of the distributions of wealth*, Quantitative Finance 4 (2004), str. 353–364.
- [31] N. Scafetta, S. Picozzi, B.J. West, *A trade-investment model for distribution of wealth*, Physica D 193 (2004), str. 338–352.
- [32] M. Patriarca, A. Chakraborti, K. Kaski, G. Germano, *Kinetic theory models for the distribution of wealth: power law from overlap of exponentials*, w książce: *Econophysics of Wealth Distributions*, pod redakcją: A. Chatterjee, B.K. Chakrabarti, S. Yarlagadda, Springer, Włochy, 2005, str. 93–110.
- [33] P. Repetowicz, S. Hutzler, P. Richmond, *Dynamics of money and income distributions*, Physica A 356 (2005), str. 641–654.
- [34] B.K. Chakrabarti, *Econophysics–Kolkata: A Short Story*, w książce: *Econophysics of Wealth Distributions*, pod redakcją: A. Chatterjee, B.K. Chakrabarti, S. Yarlagadda, Springer, Włochy, 2005, str. 225–228.

- [35] A. Chatterjee, B.K. Chakrabarti, R. B. Stinchcombe, *Master equation for a kinetic model of a trading market and its analytic solution*, Physical Review E 72 (2005), str. 026126.
- [36] P. Bhattacharyya, A. Chatterjee, B. K. Chakrabarti, *A common mode of origin of power laws in models of market and earthquake*, Physica A 381 (2005), str. 377–382.
- [37] A. Banerjee, V.M. Yakovenko, *Universal patterns of inequality*, New Journal of Physics 12 (2010), str. 075032.
- [38] S. Solomon, P. Richmond, *Power laws of wealth, market order volumes and market returns*, Physica A 299 (2001), str. 188–197.
- [39] P. Richmond, S. Solomon, *Power laws are disguised Boltzmann laws*, International Journal of Modern Physics C 12 (2001), str. 333–343.
- [40] S. Solomon, P. Richmond, *Stable power laws in variable economies; Lotka–Volterra implies Pareto–Zipf*, The European Physical Journal B 27 (2002), str. 257–261.
- [41] D.W. Huang, *Wealth accumulation with random redistribution*, Physical Review E 69 (2004), str. 057103.
- [42] N. Kiyotaki, R. Wright, *A Search–Theoretic Approach to Monetary Economics*, The American Economic Review 83 (1993), str. 63–77.
- [43] M. Molico, *The Distribution of Money and Prices in Search Equilibrium*, International Economic Review 47 (2006), str. 701–722.
- [44] M. Levy, S. Solomon, *New evidence for the power-law distribution of wealth*, Physica A 242 (1997), str. 90–94.
- [45] A. Drăgulescu, V.M. Yakovenko, *Exponential and power-law probability distributions of wealth and income in the United Kingdom and the United States*, Physica A 299 (2001), str. 213–221.
- [46] A. Drăgulescu, V.M. Yakovenko, *Evidence for the exponential distribution of income in the USA*, The European Physical Journal B 20 (2001), str. 585–589.
- [47] W.J. Reed, *The Pareto law of incomes – an explanation and an extension*, Physica A 319 (2003), str. 469–486.

- [48] P.K. Rawlings, D. Reguera, H. Reiss, *Entropic basis of the Pareto law*, Physica A 343 (2004), str. 643–652.
- [49] P. Łukasiewicz, A. Orłowski, *Probabilistic models of income distributions*, Physica A 344 (2004), str. 146–151.
- [50] V.M. Yakovenko, A.C. Silva, *Two-class Structure of Income Distribution in the USA: Exponential Bulk and Power-law Tail*, w książce: *Econophysics of Wealth Distributions*, pod redakcją: A. Chatterjee, B.K. Chakrabarti, S. Yarlagadda, Springer, Włochy, 2005, str. 15–23.
- [51] A.C. Silva, V.M. Yakovenko, *Temporal evolution of the "thermal" and "superthermal" income classes in the USA during 1983–2001*, Europhysics Letters 69 (2005), str. 304–310.
- [52] F. Clementi, M. Gallegati, *Pareto's Law of Income Distribution: Evidence for Germany, the United Kingdom, and the United States*, w książce: *Econophysics of Wealth Distributions*, pod redakcją: A. Chatterjee, B.K. Chakrabarti, S. Yarlagadda, Springer, Włochy, 2005, str. 3–14.
- [53] M. Nirei, W. Souma, *A two factor model of income distribution dynamics*, Review of Income and Wealth 53 (2007), str. 440–459.
- [54] F. Clementi, T. Di Matteo, M. Gallegati, G. Kaniadakis, *The κ -generalized distribution: A new descriptive model for the size distribution of incomes*, Physica A 387 (2008), str. 3201–3208.
- [55] F. Clementi, M. Gallegati, G. Kaniadakis, *A κ -generalized statistical mechanics approach to income analysis*, Journal of Statistical Mechanics (2009), str. P02037.
- [56] J.C. Ferrero, *The Monomodal, Polymodal, Equilibrium and Nonequilibrium Distribution of Money*, w książce: *Econophysics of Wealth Distributions*, pod redakcją: A. Chatterjee, B.K. Chakrabarti, S. Yarlagadda, Springer, Włochy, 2005, str. 159–167.
- [57] F. Clementi, M. Gallegati, G. Kaniadakis, *κ -generalized statistics in personal income distribution*, The European Physical Journal B 57 (2007), str. 187–193.

- [58] F. Clementi, M. Gallegati, *Power law tails in the Italian personal income distribution*, Physica A 350 (2005), str. 427–438.
- [59] F. Clementi, T. Di Matteo, M. Gallegati, *The power-law tail exponent of income distributions*, Physica A 370 (2006), str. 49–53.
- [60] H. Aoyama, W. Souma, Y. Nagahara, M.P. Okazaki, H. Takayasu, M. Takayasu, *Pareto's Law for Income of Individuals and Debt of Bankrupt Companies*, Fractals 8 (2000), str. 293–300.
- [61] W. Souma, *Universal Structure of the Personal Income Distribution*, Fractals 9 (2001), str. 463–470.
- [62] Y. Fujiwara, W. Souma, H. Aoyama, T. Kaizoji, M. Aoki, *Growth and fluctuations of personal income*, Physica A 321 (2003), str. 598–604.
- [63] H. Aoyama, W. Souma, Y. Fujiwara, *Growth and fluctuations of personal and company's income*, Physica A 324 (2003), str. 352–358.
- [64] W. Souma, M. Nirei, *Empirical Study and Model of Personal Income*, w książce: *Econophysics of Wealth Distributions*, pod redakcją: A. Chatterjee, B.K. Chakrabarti, S. Yarlagadda, Springer, Włochy, 2005, str. 34–42.
- [65] T. Di Matteo, T. Aste, S.T. Hyde, *Exchanges in complex networks: income and wealth distributions*, w książce: *The Physics of Complex Systems (New Advances and Perspectives)*, pod redakcją: F. Mallamace, H.E. Stanley, IOS Press, Amsterdam, 2004, str. 435–442.
- [66] A. Banerjee, V.M. Yakovenko, T. Di Matteo, *A study of the personal income distribution in Australia*, Physica A 370 (2006), str. 54–59.
- [67] S. Sinha, *Evidence for power-law tail of the wealth distribution in India*, Physica A 359 (2006), str. 555–562.
- [68] K. Kim, S.M. Yoon, *Power Law Distributions in Korean Household Incomes*, 2004, praca niepublikowana, dostępna pod adresem: <http://arxiv.org/abs/cond-mat/0403161>.

- [69] N. Derzsy, Z. Néda, M.A. Santos, *Income distribution patterns from a complete social security database*, *Physica A* 391 (2012), str. 5611–5619.
- [70] S. Galam, *Sociophysics – A Physicist’s Modeling of Psycho-political Phenomena*, Springer, New York, 2012.
- [71] M. Jagielski, R. Kutner, *Study of Households’ Income in Poland by Using the Statistical Physics Approach*, *Acta Physica Polonica A* 117 (2010), str. 615–618.
- [72] M. Jagielski, R. Kutner, *Modelowanie zamożności polskich gospodarstw domowych metodami statystycznymi*, *Ekonomia* 25 (2011), str. 154–162.
- [73] M. Jagielski, R. Kutner, M. Pęczkowski, *Preliminary Comparison of Households’ Income in Poland with European Union and United States Ones by Using the Statistical Physics Methods*, *Acta Physica Polonica A* 121 (2012), str. B47–B49.
- [74] M. Jagielski, R. Kutner, *Ab initio analysis of all income society classes in the European Union*, *Acta Physica Polonica A* 123 (2013), str. 538–541.
- [75] M. Jagielski, R. Kutner, *Modeling of income distribution in the European Union with the Fokker–Planck equation*, *Physica A* 392 (2013), str. 2130–2138.
- [76] N.G. van Kampen, *Procesy stochastyczne w fizyce i chemii*, Państwowe Wydawnictwo Naukowe, Warszawa, 1990.
- [77] K. Jacobs, *Stochastic Processes for Physicists*, Cambridge University Press, Cambridge, 1990.
- [78] T. Srokowski, *Multiplicative Lévy processes: Itô versus Stratonovich interpretation*, *Physical Review E* 80 (2009), str. 051113.
- [79] T. Srokowski, *Nonlinear stochastic equations with multiplicative Lévy noise*, *Physical Review E* 81 (2010), str. 051110.
- [80] S. Chandrasekhar, *Stochastic Problems in Physics and Astronomy*, *Review of Modern Physics* 15 (1943), str. 1–89.
- [81] R. Kubo, M. Toda, N. Hashitsume, *Fizyka statystyczna II. Mechanika statystyczna stanów nierównowagowych*, Wydawnictwo Naukowe PWN, Warszawa, 1991.

- [82] R. Kutner, *Wprowadzenie do ekonofizyki: Niegaussowskie procesy stochastyczne oraz niedebye'owska relaksacja w realu. Elementy teorii ryzyka rynkowego wraz z elementami teorii zdarzeń ekstremalnych*, 2013, skrypt dostępny na stronie http://www.fuw.edu.pl/tl_files/studia/materialy/ef/bloweier.pdf.
- [83] L.D. Landau, J.M. Lifszyc, *Fizyka statystyczna*, Wydawnictwo Naukowe PWN, Warszawa, 2012.
- [84] C.R. McConnell, S.L. Brue, S.M. Flynn, *Economics: Principles, Problems, and Policies*, McGraw-Hill, New York, 2011.
- [85] J.L. McCauley, *Nonstationarity of efficient finance markets: FX market evolution from stability to instability*, *International Review of Financial Analysis* 17 (2008), str. 820–837.
- [86] B.M. Friedman, F.H. Hahn, *Handbook of Monetary Economics*, North-Holland, Amsterdam, 1990.
- [87] R.J. Hyndman, Y. Fan, *Quantiles in Statistical Packages*, *The American Statistician* 50 (1996), str. 361–365.
- [88] O.S. Klass, O. Biham, M. Levy, O. Malcai, S. Solomon, *The Forbes 400, the Pareto power-law and efficient markets*, *The European Physical Journal B* 55 (2007), str. 143–147.
- [89] R.N. Mantegna, H.E. Stanley, *An Introduction to Econophysics: Correlations and Complexity in Finance*, Cambridge University Press, Cambridge, 2000.
- [90] A. Fronczak, P. Fronczak, *Świat sieci złożonych. Od fizyki do Internetu*, Wydawnictwo Naukowe PWN, Warszawa, 2009.
- [91] G.K. Zipf, *Human Behavior and the Principle of Least Effort*, Addison-Wesley Press, Cambridge, Massachusetts, 1949.
- [92] K. Okuyama, M. Takayasu, H. Takayasu, *Zipf's law in income distribution of companies*, *Physica A* 269 (1999), str. 125–131.

- [93] A. Baker, *Simplicity*, w książce: *The Stanford Encyclopedia of Philosophy*, pod redakcją: E.N. Zalta, Stanford University, Stanford, 2011, <http://plato.stanford.edu/archives/sum2011/entries/simplicity/>.
- [94] *Słownik pojęć Głównego Urzędu Statystycznego*, Główny Urząd Statystyczny, 2013, <http://www.stat.gov.pl/>.
- [95] V.T. Chow, D.R. Maidment, L.W. Mays, *Applied Hydrology*, McGraw-Hill, Singapore, 1998.
- [96] W.C. Haneberg, *Computational geosciences with Mathematica*, Springer, Berlin, 2004.
- [97] R. Nowak, *Statystyka dla fizyków*, Wydawnictwo Naukowe PWN, Warszawa, 2002.
- [98] E.W. Weisstein, *Levenberg-Marquardt Method*, From MathWorld – A Wolfram Web Resource, 2013, <http://mathworld.wolfram.com/Levenberg-MarquardtMethod.html>.
- [99] J. Janczura, R. Weron, *Black swans or dragon-kings? A simple test for deviations from the power law*, The European Physical Journal – Special Topics 205 (2012), str. 79–93.
- [100] K. Yamamoto, S. Miyazima, H. Yamamoto, T. Ohtsuki, A. Fujihara, *The power-law exponent and the competition rule of the high income model*, w książce: *Practical Fruits of Econophysics*, pod redakcją: H. Takayasu, Springer, Tokyo, 2006, str. 349–353.
- [101] *Rankingi Wprost*, Wprost, 2013, <http://100najbogatszych.wprost.pl/>.
- [102] S. Levy, *Wealthy People and Fat Tails: An Explanation for the Lévy Distribution of Stock Returns*, Recent Work, Finance (1998), <http://escholarship.org/uc/item/5zf0f3tg>.
- [103] M. Levy, *Are rich people smarter?*, Journal of Economic Theory 110 (2003), str. 42–64.

- [104] *Statistics on income, social inclusion and living conditions*, Eurostat, 2013, http://epp.eurostat.ec.europa.eu/portal/page/portal/income_social_inclusion_living_conditions/introduction.
- [105] EUSILC XUDB 2004 – version of August 2009.
- [106] EUSILC UDB 2005 – version 5 of August 2009.
- [107] EUSILC UDB 2006 – version 4 of March 2010.
- [108] EUSILC UDB 2007 – version 3 of March 2010.
- [109] EUSILC UDB 2008 – version 2 of August 2010.
- [110] EUSILC UDB 2009 – version 3 of March 2012.
- [111] EUSILC UDB 2010 – version 1 of March 2012.
- [112] *The World's Billionaires*, Forbes, 2013, <http://www.forbes.com/billionaires/>.
- [113] C. Gini, *On the Measure of Concentration with Special Reference to Income and Statistics*, Colorado College Publication, General Series No. 208 (1936), str. 73–79.
- [114] P.M. Dixon, J. Weiner, T. Mitchell-Olds, R. Woodley, *Bootstrapping the Gini Coefficient of Inequality*, Ecology 68 (1987), str. 1548–1551.
- [115] P.M. Dixon, J. Weiner, T. Mitchell-Olds, R. Woodley, *Bootstrapping the Gini Coefficient of Inequality – Errata*, Ecology 69 (1988), str. 1307.
- [116] C. Damgaard, J. Weiner, *Describing Inequality in Plant Size or Fecundity*, Ecology 81 (2000), str. 1139–1142.
- [117] *Gini coefficient*, Wikipedia, 2013, <http://en.wikipedia.org/>.